Random matrices

Bertrand Eynard, Taro Kimura, Sylvain Ribault

Random matrices

Bertrand Eynard, Taro Kimura and Sylvain Ribault
based on lectures by Bertrand Eynard at IPhT, Saclay
bertrand.eynard@cea.fr, taro.kimura@keio.jp, sylvain.ribault@cea.fr

October 16, 2015

Abstract

We provide a self-contained introduction to random matrices. While some applications are mentioned, our main emphasis is on three different approaches to random matrix models: the Coulomb gas method and its interpretation in terms of algebraic geometry, loop equations and their solution using topological recursion, orthogonal polynomials and their relation with integrable systems. Each approach provides its own definition of the spectral curve, a geometric object which encodes all the properties of a model. We also introduce the two peripheral subjects of counting polygonal surfaces, and computing angular integrals.

Keywords

Random matrix, eigenvalue distribution, ribbon graph, large size limit, random surface, spectral curve, loop equations, topological recursion, orthogonal polynomials, integrable systems

Contents

Chapter 0

Preliminaries

0.1 Scope

This text is based on a series of six lectures given at IPhT Saclay in early 2015, which amounted to an introduction to random matrices. Notes taken during the lectures were then completed and expanded, with the aim of introducing some of the most important approaches to random matrices. Still, this text is far from being a survey of the whole topic of random matrices, as entire areas of activity are left out. Some of the missing areas are covered in the following books or articles:

- M. L. Mehta, *Random Matrices* [1]

  The bible of random matrices, one of the first books on the topic, it covers orthogonal polynomials very well.

- P. Bleher and A. Its, eds., *Random Matrix Models and their Applications* [2]

  This book mostly deals with the large N asymptotic analysis, using the Riemann–Hilbert method.

- G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices* [3]

  This book deals with probabilistic methods, such as large deviations, concentration, and diffusion equations.

- P. J. Forrester, *Log-Gases and Random Matrices* [4]

  This book deals with the Coulomb gas approach, which leads to many interesting results, in particular formulas involving orthogonal or Jack polynomials.

- G. Akemann, J. Baik, and P. Di Francesco, eds., *The Oxford Handbook of Random Matrix Theory* [5]

  This is a topical review book, collecting contributions from many authors in random matrix theory and their applications in mathematics and physics.

- J. Harnad, ed., *Random Matrices, Random Processes and Integrable Systems* [6]

  This book focuses on the relationships of random matrices with integrable systems, fermion gases, and Grassmannians.

- V. Kazakov, M. Staudacher and T. Wynter, *Character expansion methods for matrix models of dually weighted graphs* [7]

  This article introduces the character expansion method.

- K. Efetov, *Supersymmetry in Disorder and Chaos* [8]

  This book on condensed-matter physics introduces the supersymmetric approach to random matrices.

- T. Tao and V. Vu, *Random matrices: Universality of local eigenvalue statistics up to the edge* [9]

  This article proposes a mathematical treatment of the universality of eigenvalue statistics.

The present text is intended for theoretical physicists and mathematicians. Graduate-level knowledge of linear algebra and complex analysis on the plane is assumed. Other mathematical subjects that appear, and on which some background would be welcome although not strictly necessary, are Lie groups, Riemann surfaces, and basic combinatorics.

0.2 Plan of the text

The introductory Chapter 1 explains why it is useful and interesting to study random matrices, introduces the fundamental definitions of the random matrix ensembles, and describes important universal features of random matrix models. These features are given without proofs, not only for the sake of conciseness, but also because they have several different proofs, depending on which approach is used.

The remaining Chapters are rather independent, and can be read separately.

The three central Chapters 3-5 introduce three different approaches to random matrices: the saddle point approximation with its reliance on algebraic geometry, the loop equations and the powerful topological recursion technique, and the orthogonal polynomials with their relation to integrable systems. Each approach will provide its own definition of the spectral curve – the fundamental geometrical object which characterizes a matrix model.

The two outlier Chapters are devoted to specific topics. In Chapter 2, matrix integrals are considered as formal integrals, and used for studying the combinatorics of graphs. In Chapter 6, altogether different types of matrix models are studied, where there is no invariance under conjugation.

The Bibliography is relatively short, as the cited works are selected for their potential usefulness to today's readers, leaving out many works of historical importance. The Index refers to the definitions or first occurrences of the listed terms. A few terms have several different definitions. We have often refrained from defining standard terms, in particular when adequate definitions are found in Wikipedia.

0.3 Highlighted topics

This text includes a number of ideas and results that, while not new, are often underappreciated or not as well-known as they would deserve, in particular:

- Normal matrix ensembles include both Gaussian and circular ensembles as particular cases. (See Section 1.2.2.)

- There is a fundamental distinction between convergent matrix integrals and formal matrix integrals. For the combinatorics of polygonal surfaces, we need formal integrals, not convergent integrals. (See Section 2.1.2.)

- For a given potential, the space of convergent normal matrix integrals and the space of solutions of the loop equations have the same dimension, which is given by a simple formula (4.13).
- The duality $\beta \to \frac{4}{\beta}$, which also acts nontrivially on the matrix size N, leaves the loop equations invariant. (See Section 4.1.5.)
- Loop equations can be solved perturbatively using topological recursion. (See Section 4.4.) Non-perturbative contributions to matrix integrals can also be computed systematically, and are also universal. (See Section 4.5.)
- Not only orthogonal polynomials, but also their Hilbert transforms, and the corresponding self-reproducing kernel, have expressions as minors of infinite matrices. These quantities moreover have graphical representations in terms of Motzkin paths. (See Section 5.2.)
- Random matrix models lead to integrable systems with polynomial spectral curves, which are therefore both simple and capable of approximating arbitrary curves. As prototypes of integrable systems, matrix models are therefore superior to spin chains. (See Section 5.3.)
- Itzykson–Zuber integrals obey Calogero–Moser, duality and recursion equations, and can be expressed in terms of Jack polynomials. (See Section 6.2.)

0.4 Acknowledgements

We are grateful to IPhT, Saclay for hosting the lectures, and to the members of the audience for their interest and questions. We wish to thank Gaëtan Borot for comments on the draft of this text.

Chapter 1

Introduction

1.1 Motivations, history, and applications

Matrices are everywhere in physics – they appear as soon as there are linear relations between variables. Randomness is also everywhere. So it is only natural that random matrices should arise in many places. Random matrices have however originally been introduced in order to deal with rather specific problems.

1.1.1 First appearances of random matrices in mathematics

Random matrices first appeared in statistics, when Wishart [10] generalized the chi-squared law to multivariate random data. Let X be a rectangular matrix made of p vectors of size $N \geq p$,

$$X = (X_1 \; X_2 \; \cdots \; X_p) \; , \qquad X_i = (x_{i1} \; x_{i2} \; \cdots \; x_{iN})^{\mathrm{T}} \; . \tag{1.1}$$

We are then interested in the properties of the correlation matrix

$$M = X^{\mathrm{T}} X \; , \tag{1.2}$$

which is a symmetric square matrix of size p. We assume that X is a random matrix, such that each row obeys a p-dimensional multivariate normal distribution $N_p(0, V)$ with zero mean and a covariance matrix V of size p. Then the probability law for the correlation matrix M is given by the **Wishart distribution**,

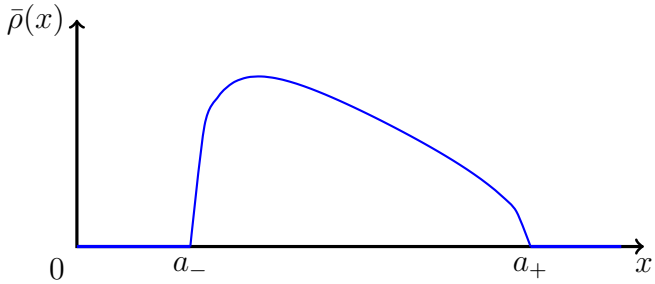
$$P(M)dM = \frac{1}{2^{\frac{Np}{2}} (\det V)^{\frac{N}{2}} \Gamma_p(\frac{N}{2})} (\det M)^{\frac{N-p-1}{2}} \, e^{-\frac{1}{2} \operatorname{Tr} V^{-1} M} dM \; . \tag{1.3}$$

After this work by Wishart on real, symmetric matrices, Ginibre studied other matrix ensembles in the 1960s, not necessarily assuming reality or even Hermiticity.

Later, in 1967, Marchenko and Pastur [11] studied the large N behaviour of the size N matrix XX^{T}, where the entries of X are independent, identically distributed variables with a normal distribution $N(0, \sigma^2)$. In the limit $N \to \infty$ with $u = \frac{p}{N} \geq 1$ fixed, the eigenvalue density was found to be given by the **Marchenko–Pastur law**,

$$\bar{\rho}(x) \propto \frac{1}{2\pi u \sigma^2} \frac{\sqrt{(a_+ - x)(x - a_-)}}{x} \; , \qquad a_\pm = \sigma^2 (1 \pm \sqrt{u})^2 \; , \tag{1.4}$$

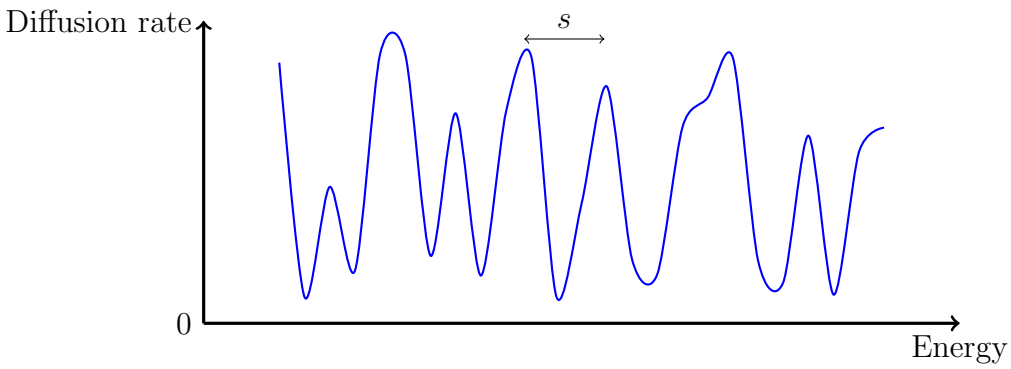
whose graph is



The eigenvalues are concentrated in a compact interval whose bounds a_-, a_+ are called the spectral edges.

1.1.2 Heavy nuclei and Wigner's idea

In the 1950s, Wigner was interested in the energy spectrums of heavy nuclei such as thorium and uranium. While theoretically much more complicated than the hydrogen nucleus, these nuclei are experimentally accessible to neutron scattering experiments. The nuclei have large numbers of energy levels, which appear in experimental data as peaks of the diffusion rate of neutrons as a function of the energy:



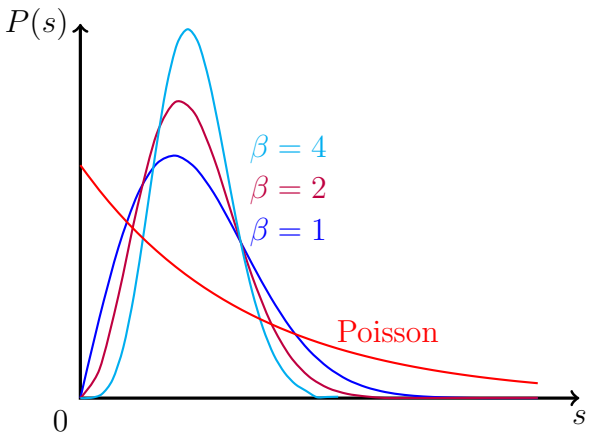
The interesting observable which Wigner studied is the statistical distribution of the distance s between neighbouring energy levels. If the energy levels were uncorrelated random numbers, the variable s would be governed by the **Poisson distribution**, with probability density

$$P(s) = e^{-s}\,. \tag{1.5}$$

But the experimentally observed probability density looks quite different, and is very well approximated (within 1%) by the **Wigner surmise**

$$P(s) = C_\beta\, s^\beta\ e^{-a_\beta s^2}\,, \tag{1.6}$$

where the parameter $\beta \in \{1, 2, 4\}$ is determined by the symmetries of the nucleus under time reversal and spin rotation, and the values of C_β and a_β are such that $\int_0^\infty ds\, P(s) = \int_0^\infty ds\, s\, P(s) = 1$.

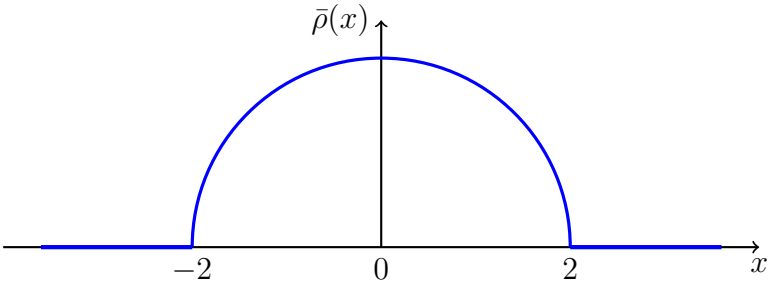


The Wigner surmise is also a very good approximation of the large size limit of the probability density for the distance between consecutive eigenvalues of random matrices. Actually, this density is universal, to the extent that it does not depend on the precise probability law of the random matrices, but only on the choice of a matrix ensemble, with real, complex and quaternionic matrices leading to $\beta = 1, 2, 4$ respectively. (See Section 1.2.) This shows that complicated Hamiltonians can be accurately modelled by random matrices.

On the other hand, the density of the eigenvalues (as opposed to their distances) is not universal. In the case of Gaussian random matrices in the limit of infinite size, it is given by the **Wigner semi-circle law**,

$$\bar{\rho}(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2} & (|x| \leq 2) \\ 0 & (|x| > 2) \end{cases} . \tag{1.7}$$

whose name comes from the shape of its graph,

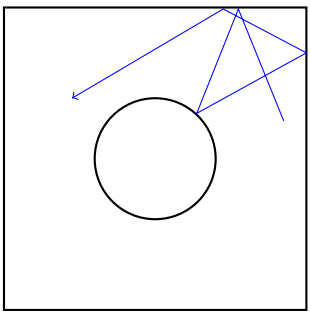


(1.8)

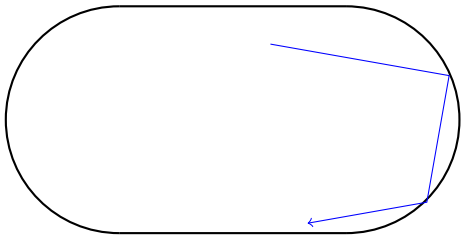
The fact that the function $\bar{\rho}(x)$ is an algebraic function and is supported on a finite interval is a feature of many random matrix models, see Chapter 3.

1.1.3 Quantum chaos in billiards

Besides heavy nuclei, other examples of systems whose Hamiltonians have very complicated spectrums are provided by quantum mechanics of a particle in 2d chaotic billiards. The particle is free, its only interaction is being reflected by the boundary. That a billiard is chaotic can already be seen at the level of classical trajectories, which are then ergodic. Billiards with rather simple shapes can be chaotic: this is the case with Sinai's billiard, and stadium-shaped billiards:

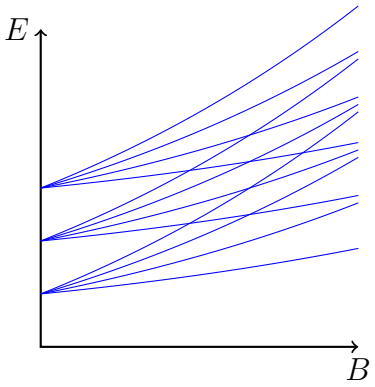


Sinai

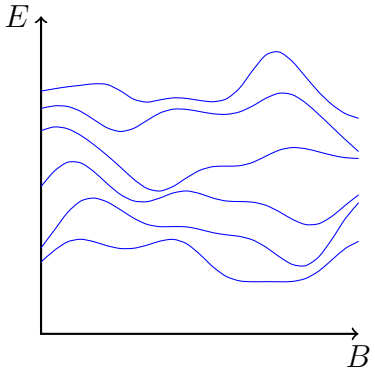


Stadium

If a billiard is classically integrable (for example a rectangle or an ellipse), then the spacings between eigenvalues of the corresponding quantum Hamiltonian follow a Poisson law, and there are many level crossings. If it is chaotic, then the eigenvalues have the same statistical properties as eigenvalues of random matrices, and they don't cross, "repelling" each other instead. This can be seen in plots of energy levels as functions of a parameter such as a magnetic field B:



Integrable billiard



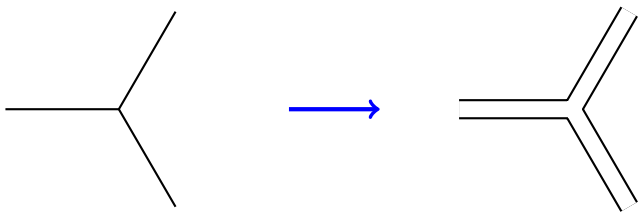
Chaotic billiard

The appearance of random matrix statistics was first observed in experiments of Berry and Tabor [12], and then raised to a major quantum chaos conjecture by Bohigas, Giannoni, and Schmit in 1984 [13]. It has been so well established experimentally that random matrix statistics are considered as experimental signatures of quantum chaos. However, this relation between quantum chaos and random matrices is still not really understood theoretically.

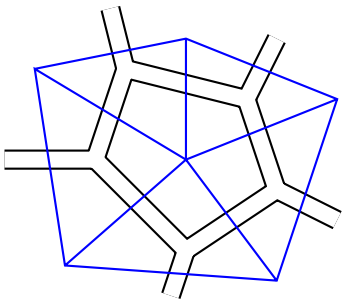
1.1.4 Quantum chromodynamics and random surfaces

Quantum chromodynamics (QCD) is a gauge theory whose gauge field, which describes gluons, has the structure of a matrix of size N_c. The integer N_c is the number of colors and is set to $N_c = 3$ in QCD. However, 't Hooft proposed to study a generalization of QCD in the limit $N_c \to \infty$ [14], where important simplifications occur, and then recover $N_c = 3$ by perturbation theory. Moreover, the $N_c \times N_c$ matrix is random, because of quantum mechanical dynamics. So, the $N_c \to \infty$ generalization of QCD is a theory of large random matrices.

In Feynman graphs of gluons, the matrix structure can be represented by having double lines in propagators, where each line implicitly carries one of the two indices of the matrix:



Such a graph with double lines is called a **ribbon graph** . In the limit $N_c \to \infty$, the graphs become planar – they can be drawn on surfaces. And a planar ribbon graph is dual to a triangulation of the corresponding surface:



So the dynamics of the surface must play a role in quantum chromodynamics.

It was initially believed that this surface dynamics would be the expected missing link between QCD and string theory, but it turned out to be more complicated. However, 't Hooft's approach was the inspiration for studying random surfaces using random matrices. This was initiated by Brézin, Itzykson, Parisi and Zuber [15], and led to the random matrix theory of 2d quantum gravity.

1.1.5 Other applications

1. **Transport in disordered systems:** Random matrices appear in disordered condensed-matter systems. In particular, random diffusion matrices can provide a microscopic explanation for the mesoscopic conductance at low temperatures, when the quantum decoherence length becomes larger than the conductor's size.

   As opposed to random Hamiltonians or random scattering matrices, such diffusion matrices live in non-compact, hyperbolic spaces where no time-invariant probability distributions exist. One can only describe their dynamics, using a kind of Fokker–Planck equation called the Dorokov–Mellow–Pereira–Kumar (DMPK) equation.

   Along similar lines, Efetov [8] has developed a supersymmetric method to deal with Gaussian random matrices, with very successful applications to solid-state physics, possibly including high-temperature superconductivity.

2. **String theory:** There are several ways to relate string theory to random matrices. All of them rather indirect. Historically, the first relation came from the identification of the string worldsheet with a random surface, through the BIPZ [15] approach. Unfortunately this approach was limited to a target space dimension $d \leq 1$, thus unable to describe string theory in 4 or 10 or 26 dimensions. Then, in 2002, Dijkgraaf and Vafa [16] observed that the tree-level prepotential in some string theories can be written in terms of a random matrix model. This had important implications in string theory, allowing many exact computations, beyond perturbation theory.

   Another relation came from Kontsevich's work [17], who used a matrix integral, now known as Kontsevich's integral, to prove Witten's conjecture on 2d quantum gravity.

Recently, other relations were found between topological string theories and matrix models. In particular, in 2008, the Bouchard–Klemm–Mariño–Pasquetti (BKMP) conjecture [18] (since then proved [19]) claimed that topological string amplitudes obey the same recursion relations as matrix models' correlation functions. These recursion relations will appear in Chapter 4 under the name of topological recursion relations.

3. **Knot theory:** Knots are characterized by topological invariants such as Jones polynomials, Alexander polynomials, and HOMFLY polynomials. Like the topological string amplitudes of the BKMP conjecture, these invariants are found to obey the same recursion relations as matrix integrals. In the special case of torus knots, the HOMFLY polynomials are actually expressed as matrix integrals. We shall mention some of this in Section 3.4.1.

4. **Conformal field theory in two dimensions:** Penner-type random matrix integrals [20] are formally identical to the Dotsenko–Fateev integrals of two-dimensional conformal field theory [21, 22]. This technical agreement has no deep explanation so far, but it implies that random matrix techniques can compute certain CFT correlation functions. This is especially useful in heavy asymptotic limits.

5. **Integrable systems:** In a random matrix model, the eigenvalue distribution is very often related to the tau function of some integrable system. The tau function in question plays the role of the partition function of the system, and obeys important relations such as Hirota equations or Sato relations.

   For example, the level spacing distribution of the Gaussian Unitary Ensemble is the tau function of the Painlevé V integrable system, also known as the Fredholm determinant of the sine kernel. The Tracy–Widom law is the tau function of the Painlevé II integrable system, also known as the Fredholm determinant of the Airy kernel.

   Random matrix models are good prototypes of integrable systems, in the sense that many relations which can be derived in random matrix theory, actually extend to all integrable systems. This general principle has been extremely fruitful so far. This is a strong motivation for studying universal structures in random matrix models. See Chapter 5.

6. **Crystal growth:** A simple model of a crystal is a 3d partition, i.e. a pile of small cubic boxes. It has been experimentally and numerically observed that the statistics of cubes on a growing crystal obey random matrix laws. In fact, there is a matrix integral which coincides with the partition function of uniformly distributed 3d partitions in a box [23].

7. **Hele-Shaw cell:** This is a model for the growth of a pocket of a non-viscous liquid (such as water) in a highly viscous liquid (such as oil) in two dimensions. The problem is to predict the evolution of the shape of the pocket, knowing that the high viscosity keeps all its moments except the area fixed. Unexpectedly, the Hele-Shaw problem was solved using random matrices [24, 25], and it was found that the pocket has the shape of the large size spectrum (thus a domain of $\mathbb{C}$) of a complex random matrix. (See Section 3.4.3.)

8. **Number theory:** In 1973, in a discussion between Montgomery and Dyson [26], it was observed numerically that the distribution of the non-trivial zeros of the Riemann zeta function on the axis $\Re(s) = \frac{1}{2}$ agrees extraordinarily well with the

eigenvalues distribution in a Circular Unitary Ensemble – a random matrix ensemble [27]. A proof of this observation would not only imply the Riemann hypothesis, but also have deep implications for the statistical distribution of prime numbers. This is considered a very promising approach in number theory.

9. **Telecommunication:** Telecommunication signals can be modelled as random vectors, and a group of signals arriving at an antenna forms a large random matrix. Real-time treatment of signals may mean inverting or diagonalizing such huge matrices. Random matrix theory is helpful for addressing this challenge. Moreover, one way to optimize transmission is to ensure that signal packages repel one another, in the same way as random matrix eigenvalues repel one another.

10. **Biology, etc:** Ribbon graphs are a simple model of RNA's secondary structure. Random matrix models, which generate ribbon graphs, can therefore help predict the shape of RNA. Some of the best RNA shape codes were developed using random matrix techniques [28].

    Moreover, random matrix statistics were observed to describe the distributions of synapses in neurons, trees in a rainforest, waiting times between buses, prices in a stock exchange, correlations in large random networks, etc.

1.2 Random matrix laws

A random matrix is a matrix whose elements are randomly distributed. A random matrix model is characterized by a matrix ensemble E, and a (not necessarily normalized) probability measure $d\mu(M)$ for $M \in E$, called the random matrix law. The **partition function** or matrix integral is then defined as

$$\boxed{\mathcal{Z} = \int_E d\mu(M)} \,. \tag{1.9}$$

1.2.1 Gaussian and circular ensembles

Matrix ensembles

The three matrix ensembles whose eigenvalue spacing distributions are approximated by the Wigner surmise are called the Gaussian ensembles or Wigner ensembles:

- the Gaussian Orthogonal Ensemble **GOE** of real symmetric matrices,
- the Gaussian Unitary Ensemble **GUE** of complex Hermitian matrices,
- and the Gaussian Symplectic Ensemble **GSE** of quaternionic Hermitian matrices.

Eigenvalue spacing distributions are universal, and do not depend on probability measures. So they were originally studied using simple Gaussian measures, and the ensembles were themselves called Gaussian. This is however a potentially confusing abuse of language, as these ensembles are also considered with non-Gaussian measures.

A matrix M in one of Wigner's three Gaussian ensembles has real eigenvalues, and can be diagonalized as

$$\boxed{M = U\Lambda U^{-1} \quad \text{with} \quad \Lambda = \text{diag}(\lambda_1, \cdots \lambda_N) \in \mathbb{R}^N} \,. \tag{1.10}$$

This is called the **angular-radial decomposition**, where Λ is the radial and U the angular part. The matrix U belongs to a compact Lie group, which we call the corresponding circular ensemble:

- the Circular Orthogonal Ensemble **COE** of real orthogonal matrices,
- the Circular Unitary Ensemble **CUE** of complex unitary matrices,
- the Circular Symplectic Ensemble **CSE** of complex symplectic matrices.

We will denote the Gaussian ensembles as E_N^β, and the corresponding circular ensembles as U_N^β, with $\beta \in \{1, 2, 4\}$:

β	Ensemble type	Gaussian ensemble	E_N^β	Circular ensemble	U_N^β
1	orthogonal	GOE	S_N	COE	O_N
2	unitary	GUE	H_N	CUE	U_N
4	symplectic	GSE	Q_N	CSE	Sp_{2N}

(1.11)

Each one of these ensembles can be realized as a space of square matrices of size N, whose coefficients are real if $\beta = 1$, complex if $\beta = 2$, and **quaternionic** if $\beta = 4$. In all cases, the coefficients can be written as elements of a β-dimensional **Clifford algebra** over $\mathbb{R}$ with generators $(\mathbf{e}_\alpha) = (\mathbf{e}_0 = \mathbf{1}, \mathbf{e}_i)$ and relations

$$\mathbf{e}_i^2 = -\mathbf{1} \ , \quad \mathbf{e}_i \mathbf{e}_j = \epsilon_{ijk} \mathbf{e}_k \ . \tag{1.12}$$

The Clifford algebra also comes with the conjugation

$$q = \sum_{\alpha=0}^{\beta} q^{(\alpha)} \mathbf{e}_\alpha \quad \Rightarrow \quad \bar{q} = q^{(0)} \mathbf{1} - \sum_{i=1}^{\beta-1} q^{(i)} \mathbf{e}_i \quad \text{where} \quad q^{(\alpha)} \in \mathbb{R} \ . \tag{1.13}$$

Defining the conjugate of a matrix by $M^\dagger = \bar{M}^T$, Gaussian ensembles are then defined by the constraint $M^\dagger = M$, and circular ensembles by the orthonormality constraint $MM^\dagger = \mathrm{Id}$. The two quaternionic ensembles have alternative realizations in terms of complex matrices of size $2N$, which are obtained by replacing the generators $\mathbf{e}_i$ with size two Pauli matrices.

In the diagonalization (1.10) of $M \in E_N^\beta$, several different choices of $U \in U_N^\beta$ are possible. In the case $\beta = 2$, the set $\{\lambda_1, \cdots \lambda_N\}$ of eigenvalues is left unchanged if we multiply U on the right by elements of

- the set $(U_1)^N \subset U_N$ of diagonal unitary matrices,
- and the set $\mathfrak{S}_N \subset U_N$ of permutation matrices.

Therefore,

$$H_N \simeq \frac{\frac{U_N}{(U_1)^N} \times \mathbb{R}^N}{\mathfrak{S}_N} \ , \tag{1.14}$$

which is a special case of

$$E_N^\beta \simeq \frac{\frac{U_N^\beta}{T_N^\beta} \times \mathbb{R}^N}{\mathrm{Weyl}(U_N^\beta)} \ , \tag{1.15}$$

where T_N^β is the maximal torus of U_N^β, and $\mathrm{Weyl}(U_N^\beta)$ is its Weyl group. Strictly speaking, this isomorphism only holds on the subset of E_N^β where the eigenvalues are distinct. Since the complement of this subset has measure zero, this restriction does not matter for matrix integrals.

Probability measures

On each one of the three Gaussian ensembles E_N^β, there is a **Lebesgue measure**, which is the product of the Lebesgue measures on the real components of the matrix M. Writing the real components of $M_{i,j}$ as $M_{i,j}^{(\alpha)}$, the Lebesgue measure is

$$dM = \prod_i dM_{i,i} \prod_{i<j} \prod_{\alpha=0}^{\beta-1} dM_{i,j}^{(\alpha)} \ . \tag{1.16}$$

The Lebesgue measure is invariant under changes of bases. Interesting measures are usually built from dM by multiplication with a function of M. For example, given a function V called the **potential**, we can consider the measure

$$\boxed{d\mu(M) = e^{-\operatorname{Tr} V(M)} dM} \ . \tag{1.17}$$

The potential V is often chosen to be a polynomial. Gaussian statistics are obtained when V is a polynomial of degree two. Generalizing polynomial potentials, we can consider **rational potentials**, which we define as potentials V such that V' is a rational function. (In the literature, rational potentials are sometimes called semi-classical potentials.) Far more exotic potentials can also be considered.

The basic measure on a circular ensemble is the **Haar measure**, which is characterized as being invariant under the left and right actions of the corresponding Lie group on itself. For example, viewing U_N as a submanifold of the set $M_N(\mathbb{C})$ of size N complex matrices, the Haar measure dM_{Haar} on U_N can be written in terms of the Lebesgue measure dM on $M_N(\mathbb{C})$ as

$$dM_{\text{Haar}} = \delta(MM^\dagger - \text{Id})\, dM \ , \tag{1.18}$$

which involves the Dirac delta function on $M_N(\mathbb{C})$.

Under diagonalization (1.10), the Lebesgue measure dM can be rewritten in terms of measures on Λ and U. As we will shortly show, we have

$$\boxed{dM = |\Delta(\Lambda)|^\beta \ d\Lambda \, dU_{\text{Haar}}} \ , \tag{1.19}$$

where $d\Lambda = \prod_{i=1}^N d\lambda_i$ is the Lebesgue measure on $\mathbb{R}^N$, and the Jacobian is written in terms of the **Vandermonde determinant**,

$$\boxed{\Delta(\Lambda) = \prod_{1\leq i<j\leq N} (\lambda_i - \lambda_j)} \ . \tag{1.20}$$

The partition function associated to the measure $d\mu(M)$ (1.17) becomes

$$\boxed{\mathcal{Z} = \text{Vol}\left(\frac{E_N^\beta}{\mathbb{R}^N}\right) \int_{\mathbb{R}^N} d\Lambda \ |\Delta(\Lambda)|^\beta \, e^{-\operatorname{Tr} V(\Lambda)}} \ , \tag{1.21}$$

where the prefactor is

$$\text{Vol}\left(\frac{E_N^\beta}{\mathbb{R}^N}\right) = \frac{1}{\left|\text{Weyl}(U_N^\beta)\right|} \text{Vol}\left(\frac{U_N^\beta}{T_N^\beta}\right) \ . \tag{1.22}$$

Explicitly,

$$\mathrm{Vol}\left(\frac{E_N^\beta}{\mathbb{R}^N}\right) \overset{\beta=1}{=} \frac{2^N \pi^{\frac{N(N+1)}{4}}}{N! \prod_{j=1}^N \Gamma(\frac{j}{2})} , \tag{1.23}$$

$$\overset{\beta=2}{=} \frac{\pi^{\frac{N(N-1)}{2}}}{N! \prod_{j=0}^{N-1} j!} , \tag{1.24}$$

$$\overset{\beta=4}{=} \frac{\pi^{N(N-1)}}{N! \prod_{j=0}^{N-1} (2j+1)!} . \tag{1.25}$$

If however the measure $d\mu(M)$ is not invariant under conjugations, then the partition function does not so easily reduce to an integral over eigenvalues. See Chapter 6 for examples of this more complicated case.

Let us now prove eq. (1.19). The Lebesgue measure dM is invariant under conjugation of M by a circular matrix, and the Haar measure dU_{Haar} is invariant under right and left multiplication, therefore the Jacobian $\frac{dM}{d\Lambda dU_{\text{Haar}}}$ in eq. (1.19) must be invariant under U_N^β conjugations, and can without loss of generality be computed at $U = \mathrm{Id}$. Denoting δM the differential of M, we have

$$(\delta M)_{ij} = \Big(\delta\Lambda + [\delta U, \Lambda]\Big)_{ij} = \delta_{ij}\delta\lambda_i + (\lambda_i - \lambda_j)\delta U_{ij} . \tag{1.26}$$

This leads to the Jacobian

$$\frac{dM}{d\Lambda dU_{\text{Haar}}} = \det_{i,i'}\left(\frac{\delta M_{ii}}{\delta\lambda_{i'}}\right) \det_{i\neq j, i'\neq j'}\left(\frac{\delta M_{ij}}{\delta U_{i'j'}}\right) = |\Delta(\Lambda)|^\beta . \tag{1.27}$$

1.2.2 Normal matrices

For $\beta = 2$ and $\gamma \subset \mathbb{C}$, we define the corresponding ensemble of **normal matrices** as

$$\boxed{H_N(\gamma) = \Big\{M = U\Lambda U^{-1} \Big| U \in U_N,\ \Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_N),\ \lambda_i \in \gamma\Big\}} . \tag{1.28}$$

This is inspired by the characterization (1.10) of Hermitian matrices, where we replaced the condition that the eigenvalues are real, with the condition that they belong to γ. Actually, normal matrices generalize not only Hermitian, but also unitary matrices:

$$H_N(\mathbb{R}) = H_N , \tag{1.29}$$

$$H_N(S^1) = U_N . \tag{1.30}$$

Normal matrices are in particular useful in the saddle point approximation (see Chapter 3). In this approximation, one should integrate the eigenvalues on steepest descent contours γ passing through a saddle point. Very often, the saddle point is complex, and the contour γ is not the real line, but another open Jordan arc going from ∞ to ∞.

Generalizing our results on Hermitian matrices, we have

$$H_N(\gamma) \simeq \frac{\frac{U_N}{(U_1)^N} \times \gamma^N}{\mathfrak{S}_N} , \tag{1.31}$$

and we define the natural measure on $H_N(\gamma)$ as

$$dM = \Delta(\Lambda)^2 d\Lambda\, dU_{\text{Haar}} , \tag{1.32}$$

where $d\Lambda = \prod_{i=1}^N d\lambda_i$ is built from the curvilinear measure on γ. For any parametrization $f : [0,1] \mapsto \mathbb{C}$ of $\gamma = f([0,1])$, that curvilinear measure can be written as

$$d\lambda_i = f'(t_i)\, dt_i \ . \tag{1.33}$$

The measure dM is in general not positive. And the partition function (1.21) becomes

$$\boxed{\mathcal{Z}(\gamma^N) = \mathrm{Vol}\left(\frac{H_N}{\mathbb{R}^N}\right) \int_{\gamma^N} d\Lambda\, \Delta(\Lambda)^2 e^{-\operatorname{Tr} V(\Lambda)}} \ . \tag{1.34}$$

In contrast to the case of the Gaussian ensembles, the factor $\Delta(\Lambda)^2$ does not come with an absolute value or modulus. This makes the integrand analytic (assuming the potential is analytic), and the partition function is invariant under small deformations of the integration domain γ^N.

More generally, for any complex value of β, we define the partition function

$$\boxed{\mathcal{Z}(\Gamma) \propto \int_\Gamma d\Lambda\, \Delta(\Lambda)^\beta e^{-\operatorname{Tr} V(\Lambda)}} \ , \tag{1.35}$$

for an integration domain $\Gamma \subset \mathbb{C}^N$ – cuts, such that the integral is absolutely convergent. The presence of cuts is necessary for $\Delta(\Lambda)^\beta$ to be well-defined, in particular there must be cuts ending at the hyperplanes $\{\lambda_i = \lambda_j\}$. Due to the analyticity of the integrand, the partition function is unchanged under homotopic deformations of Γ, and therefore only depends on Γ through its homology class. For rational values of β, the homology space of Γ is finite-dimensional.

Integration contours

We have just seen the interest of integrating matrix eigenvalues on fairly arbitrary contours in the complex plane. Let us now determine which contours give rise to convergent integrals.

For a rational potential V, the convergence of the partition function $\mathcal{Z}(\gamma^N)$ is equivalent to the convergence of the integral

$$z(\gamma) = \int_\gamma e^{-V(\lambda)} d\lambda \ . \tag{1.36}$$

Let us study this integral near a singularity α of V': either $\alpha = \infty$, or α is a pole of V'. We define the degrees of V and V' at their singularities by

$$\deg_\alpha \frac{1}{(\lambda-\alpha)^n} = n \quad , \quad \deg_\alpha \log(\lambda - \alpha) = 0 \quad , \quad \deg_\infty \lambda^n = n \ . \tag{1.37}$$

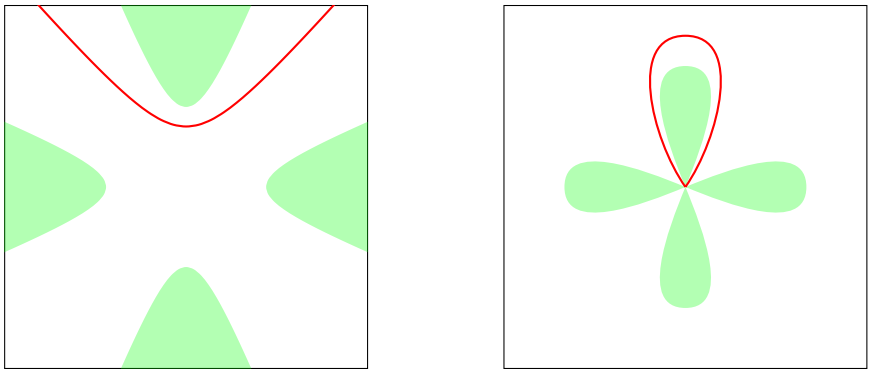
The total **degree** of V' is

$$\boxed{d = \deg V' = \sum_\alpha \deg_\alpha V'} \ , \tag{1.38}$$

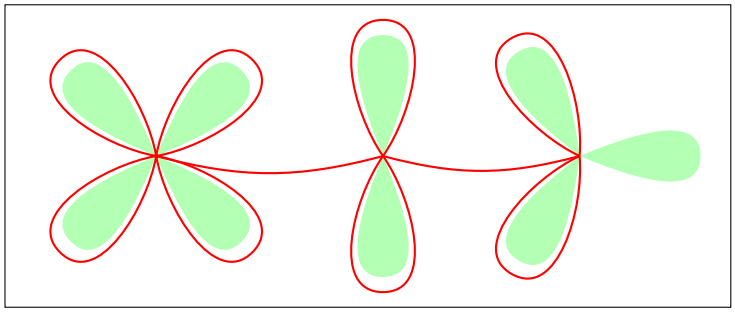
and V' has $d+1$ independent parameters. We also define the corresponding **divisor** as the formal combination of singularities,

$$\operatorname{div} V' = \sum_\alpha (\deg_\alpha V') \cdot [\alpha] \ . \tag{1.39}$$

An integration contour γ can end at a singularity α that is not a simple pole, provided $\lim_{\lambda\in\gamma,\lambda\to\alpha} \Re V(\lambda) = +\infty$. This defines $\deg_\alpha V$ allowed sectors where $\Re V(\lambda) \to +\infty$, and $\deg_\alpha V$ forbidden sectors where $\Re V(\lambda) \to -\infty$. For example, in the case $\deg_\alpha V = 4$ with $\alpha = \infty$ (left) or $\alpha \neq \infty$ (right), we draw the four forbidden sectors in green, and an example of an allowed integration contour in red:



When it comes to finding integration contours, these two examples are actually equivalent, and it does not matter whether $\alpha = \infty$ or not. In a case with three singularities such that $\deg_\alpha V = 2, 3, 4$, let us draw a set of topologically independent contours:



In general, each extra singularity α gives rise to $\deg_\alpha V + 1$ extra independent contours. However, the first singularity gives rise to only $\deg_\alpha V - 1$ independent contours. If by convention we consider that first singularity to be $\alpha = \infty$, then we have $\deg_\alpha V'$ independent contours per singularity, and the total number of independent contours is the total degree d of V'. In other words, if we define $\mathbb{C}_{\text{allowed}}$ as the complex λ-plane minus the forbidden sectors, the rank of the first fundamental group of $\mathbb{C}_{\text{allowed}}$ is

$$\operatorname{rank} \pi_1(\mathbb{C}_{\text{allowed}}) = d \ . \tag{1.40}$$

Arbitrary linear combinations of these contours form the d-dimensional **homology space**

$$H_1(e^{-V(\lambda)}d\lambda) = \left\{ \gamma = \sum_{i=1}^{d} c_i \gamma_i, \ c_i \in \mathbb{C} \right\} \ . \tag{1.41}$$

These results also hold if V' has simple poles. For α a simple pole of V', the integrand behaves as $e^{-V(\lambda)} \underset{\lambda\to\alpha}{\propto} (\lambda - \alpha)^{-t_\alpha}$ with $t_\alpha = \operatorname{Res}_\alpha V'$. A simple pole always gives rise to $\deg_\alpha V' = 1$ extra contour: either around α if $t_\alpha > 0$, or ending at α if $t_\alpha < 0$. (The condition for integrability would actually be $t_\alpha < 1$, but we need $e^{-V(\alpha)} = 0$ for integrating by parts.) Possible branch cuts do not modify the counting of contours. There is a branch cut ending at α if $t_\alpha \notin \mathbb{Z}$. A branch cut, if present, must end at another singularity, possibly at infinity: this is always possible, because $\sum_\alpha t_\alpha = 0$.

Filling fractions

The normal matrix ensemble $H_N(\gamma)$, and the corresponding convergent matrix integral $\mathcal{Z}(\gamma^N)$, can be defined not only for individual contours, but also for their linear combinations $\gamma \in H_1(e^{-V(\lambda)}d\lambda)$. We now relax the requirement that the N-dimensional integration domain be symmetric under permutations, and consider the contours

$$\vec{\gamma}^{\vec{n}} = \prod_{i=1}^{d} \gamma_i^{n_i} \quad \text{with} \quad \sum_{i=1}^{d} n_i = N \ , \tag{1.42}$$

where n_i eigenvalues are integrated over the contour γ_i. The quantities

$$\epsilon_i = \frac{n_i}{N} \tag{1.43}$$

are then called the **filling fractions** associated with our basis γ_i of contours. (Changes of basis $\gamma_i \to \sum_{i,j} C_{i,j}\gamma_j$ act linearly on filling fractions, $\epsilon_i \to \sum_{i,j} C_{i,j}\epsilon_j$.) The corresponding integral is called $\mathcal{Z}(\vec{\gamma}^{\vec{n}})$. Linearly combining N-dimensional contours, we define

$$\mathcal{Z}\left(\sum_{\vec{n}} c_{\vec{n}}\vec{\gamma}^{\vec{n}}\right) = \sum_{\vec{n}} c_{\vec{n}}\mathcal{Z}(\vec{\gamma}^{\vec{n}}) \ , \quad \text{with } c_{\vec{n}} \in \mathbb{C} \ . \tag{1.44}$$

The dimension of the space of convergent integrals, equivalently of the homology space of the one-form $\prod_{i=1}^{N} e^{-V(\lambda_i)}d\lambda_i$, is therefore the number of d-uples $\vec{n}$ whose sum is N,

$$\boxed{d_N = \dim H_1\left(\prod_{i=1}^{N} e^{-V(\lambda_i)}d\lambda_i\right) = \binom{N+d-1}{N}} \ . \tag{1.45}$$

For any contour $\gamma = \sum_{i=1}^{d} c_i\gamma_i \in H_1(e^{-V(\lambda)}d\lambda)$, the normal matrix integral $\mathcal{Z}(\gamma^N)$ can be decomposed as

$$\frac{1}{N!}\mathcal{Z}(\gamma^N) = \sum_{\vec{n}} \frac{\prod_{i=1}^{n} c_i^{n_i}}{\prod_{i=1}^{d} n_i!}\mathcal{Z}(\vec{\gamma}^{\vec{n}}) \ . \tag{1.46}$$

Unitary matrices as normal matrices

We now know two measures on U_N: the original Haar measure dM_{Haar}, and the measure dM which follows from the realization of $U_N = H_N(S^1)$ as a normal matrix ensemble. As we will shortly show, the relation between these measures is

$$dM_{\text{Haar}} = (-1)^{\frac{N(N-1)}{2}}(i\det M)^{-N}dM = (-1)^{\frac{N(N-1)}{2}}i^{-N}e^{-N\,\text{Tr}\log M}dM \ . \tag{1.47}$$

This relates the positive but apparently non-analytic Haar measure, with the analytic but not positive measure dM. This is useful, because an integral with dM can be computed by deforming integration contours. It can actually happen that the only poles come from the prefactor $(\det M)^{-N}$, in which case the integral reduces to a sum of residues at $\lambda_i = 0$.

Let us now derive the relation between the Haar and normal measures. We diagonalize our matrix $M \in U_N$ as $M = U\Lambda U^\dagger$, and we parametrize the eigenvalues $\lambda_i \in S^1$ by $\lambda_i = e^{i\theta_i}$ with $\theta_i \in [0, 2\pi]$. Let us compute the Haar measure at $U = \text{Id}$, by first computing the invariant metric which induces that measure:

$$\begin{aligned} \text{Tr}(\delta M\delta M^\dagger) &= \text{Tr}\left(\delta\Lambda + [\delta U, \Lambda]\right)\left(\delta\Lambda + [\delta U, \Lambda]\right)^\dagger \ , &(1.48)\\ &= \text{Tr}\left(\delta\Lambda\delta\Lambda^\dagger + [\delta U, \Lambda][\delta U, \Lambda]^\dagger\right) \ , &(1.49)\\ &= \sum_{i=1}^{N}\delta\theta_i^2 + 2\sum_{i<j}|\lambda_i - \lambda_j|^2|\delta U_{ij}|^2 \ . &(1.50) \end{aligned}$$

The Haar measure is therefore

$$dM_{\text{Haar}} = dU_{\text{Haar}} \prod_{i<j} |\lambda_i - \lambda_j|^2 \prod_{i=1}^{N} d\theta_i \ . \tag{1.51}$$

On the other hand, the measure dM (1.32) is

$$dM = dU_{\text{Haar}} \prod_{i<j} (\lambda_i - \lambda_j)^2 \prod_{k=1}^{N} d\lambda_k \ , \tag{1.52}$$

$$= dU_{\text{Haar}} \prod_{i<j} (-\lambda_i \lambda_j) |\lambda_i - \lambda_j|^2 \prod_{k=1}^{N} i\lambda_k d\theta_k \ , \tag{1.53}$$

$$= (-1)^{\frac{N(N-1)}{2}} \left(i \prod_{k=1}^{N} \lambda_k \right)^N dM_{\text{Haar}} \ . \tag{1.54}$$

1.2.3 Other examples

1. **Complex matrix model:** The ensemble of complex matrices is $M_N(\mathbb{C})$, with its Lebesgue measure dM. Let us consider a matrix integral of the form

$$\boxed{\mathcal{Z} = \int_{M_N(\mathbb{C})} dM \, e^{-\operatorname{Tr} V(M, M^\dagger)}} \ . \tag{1.55}$$

This is invariant under conjugations of M by unitary matrices. Using such conjugations, we cannot diagonalize M, but only bring it to a triangular form

$$M = U(\Lambda + T)U^{-1} \ , \quad \begin{cases} U \in U_N \ , \\ \Lambda \in \mathbb{C}^N \ , \\ T \in B_N(\mathbb{C}) \ , \end{cases} \tag{1.56}$$

where $B_N(\mathbb{C})$ is set of strictly upper triangular matrices. This amounts to writing our complex matrix ensemble as

$$M_N(\mathbb{C}) \simeq \frac{\frac{U_N}{(U_1)^N} \times \mathbb{C}^N \times B_N(\mathbb{C})}{\mathfrak{S}_N} \ . \tag{1.57}$$

Under this decomposition, it can be shown that the Lebesgue measure on $M_N(\mathbb{C})$ becomes

$$\boxed{dM = |\Delta(\Lambda)|^2 |d\Lambda|^2 |dT|^2 dU_{\text{Haar}}} \ , \tag{1.58}$$

where $|d\Lambda|^2$ and $|dT|^2$ are the Lebesgue measures on $\mathbb{C}^N$ and $B_N(\mathbb{C})$.

Let us now specialize to the Gaussian complex matrix model,

$$\operatorname{Tr} V(M) = \operatorname{Tr} MM^\dagger = \operatorname{Tr} \left(\Lambda \bar{\Lambda} + TT^\dagger \right) \ . \tag{1.59}$$

Let us show how the calculation of moments can be reduced to integrals over eigenvalues. This is straightforward for moments which only involve eigenvalues, such as

$$\left\langle \operatorname{Tr} M^k \right\rangle = \frac{1}{\mathcal{Z}} \int_{M_N(\mathbb{C})} dM \ e^{-\operatorname{Tr} MM^\dagger} \operatorname{Tr} M^k \ , \tag{1.60}$$

$$= \frac{\int |d\Lambda|^2 \prod_{i<j} |\lambda_i - \lambda_j|^2 \prod_i e^{-|\lambda_i|^2} \left(\sum_i \lambda_i^k \right)}{\int |d\Lambda|^2 \prod_{i<j} |\lambda_i - \lambda_j|^2 \prod_i e^{-|\lambda_i|^2}} \ . \tag{1.61}$$

This is more complicated for moments such as $\left\langle \operatorname{Tr} M^2 M^{\dagger 2} \right\rangle$, which also involve the triangular matrix T:

$$\operatorname{Tr} M^2 M^{\dagger 2} = \operatorname{Tr}(\Lambda + T)^2 (\bar{\Lambda} + T^\dagger)^2 \ , \tag{1.62}$$

$$= \operatorname{Tr}\left(\Lambda^2 \bar{\Lambda}^2 + \Lambda \bar{\Lambda} T T^\dagger + \Lambda \bar{\Lambda} T^\dagger T + \bar{\Lambda} T \Lambda T^\dagger + \Lambda T \bar{\Lambda} T^\dagger + T^2 T^{\dagger 2} \right) \ . \tag{1.63}$$

However, the integral over T is Gaussian, and can be performed using Wick's theorem:

$$\int |dT|^2 e^{-\operatorname{Tr} T T^\dagger} \operatorname{Tr}\left(\Lambda^2 \bar{\Lambda}^2 + \Lambda \bar{\Lambda} T T^\dagger + \Lambda \bar{\Lambda} T^\dagger T + \bar{\Lambda} T \Lambda T^\dagger + \Lambda T \bar{\Lambda} T^\dagger + T^2 T^{\dagger 2} \right)$$

$$= \left(\int |dT|^2 e^{-\operatorname{Tr} T T^\dagger} \right) \left(\sum_i |\lambda_i|^4 + (N-1) \sum_i |\lambda_i|^2 + \sum_{i \neq j} \lambda_i \bar{\lambda}_j + \binom{N}{3} \right) \ . \tag{1.64}$$

2. **Multi-matrix models:** The matrix models which we have considered so far could be called **one-matrix models**, as the corresponding integrals involved only one matrix. A natural generalization is to consider integrals over multiple matrices, and the corresponding multi-matrix models. For example, a **two-matrix model** can be defined from the ensemble $E = H_N \times H_N$ and the measure

$$\boxed{d\mu(M_1, M_2) = e^{-\operatorname{Tr}(V_1(M_1) + V_2(M_2) - M_1 M_2)} dM_1 dM_2} \ , \tag{1.65}$$

where V_1 and V_2 are two functions. This can be generalized to the **matrix chain** on $(H_N)^k$,

$$\boxed{d\mu(M_1, \ldots, M_k) = e^{-\operatorname{Tr} \sum_{i=1}^k V_i(M_i)} e^{\operatorname{Tr} \sum_{i=1}^k M_i M_{i+1}} \prod_{i=1}^k dM_i} \ , \tag{1.66}$$

where the fixed matrix M_{k+1} is called an **external field**, and breaks the invariance under conjugation. See Chapter 6 for calculations in the presence of an external field in the case $k = 1$.

3. **Alternative probability laws** for Hermitian matrices include

$$d\mu(M) = e^{-\operatorname{Tr} V(M)} \det(\mathrm{Id} \otimes M + M \otimes \mathrm{Id})^{-\frac{n}{2}} dM \ , \tag{1.67}$$

which defines the $O(n)$ **matrix model**, and

$$d\mu(M) = e^{-\operatorname{Tr} V(M)} \det(\Gamma(M)) dM \ , \tag{1.68}$$

which appears in crystal growth and string theory. Here $\Gamma(M)$ is Euler's Gamma function, or some rational combination built from it.

4. **Matrix models for knot theory**: Given two relatively prime integers (P, Q), let $\mathfrak{K}_{P,Q}$ be the knot which winds P times around the meridian and Q times around the longitude of a two-dimensional torus. For example, $\mathfrak{K}_{3,2}$ is the trefoil knot, and $\mathfrak{K}_{1,1}$

is the unknot with the standard framing. Let us consider the matrix model with the ensemble H_N and the measure

$$d\mu(M) = e^{-\frac{\operatorname{Tr} M^2}{2PQ\log q}} \det\left(f_P(\mathrm{Id}\otimes M, M\otimes \mathrm{Id})\right) \det\left(f_Q(\mathrm{Id}\otimes M, M\otimes \mathrm{Id})\right) dM \ , \tag{1.69}$$

where

$$f_P(x,x') = \frac{\sinh\frac{x-x'}{2P}}{x-x'} \ . \tag{1.70}$$

Then the expectation values of Schur polynomials s_R of the eigenvalues λ_i of our random matrix turn out to be the colored HOMFLY polynomials of the torus knot $\mathfrak{K}_{P,Q}$ [29],

$$\left\langle s_R(e^{\lambda_1},\dots,e^{\lambda_N})\right\rangle = \mathrm{HOMFLY}_R(\mathfrak{K}_{P,Q}, SU_N, q) \ , \tag{1.71}$$

where the partition R also serves as the label of a representation of SU_N. For $N=2$, colored HOMFLY polynomials reduce to colored Jones polynomials. The invariants of more general knots can also be written as Vassiliev–Kontsevich matrix integrals, which become substantially more complicated when the knots are not torus knots.

5. **Algebraic submanifolds** of $M_N(\mathbb{C})$ can be used as matrix ensembles. This generalizes normal matrix ensembles (including the unitary ensemble), which are algebraic submanifolds of $M_N(\mathbb{C})$ because conditions on eigenvalues can be written in terms of the characteristic polynomial.

6. **Supermatrix models** are built from matrices with fermionic entries. The relevant ensemble is a supermanifold, for example the ensemble of super-Hermitian matrices, and the corresponding circular ensembles are supergroups.

Other examples can be built by combining or generalizing the above ensembles and measures.

Classification by involutions: symmetric spaces

There were many attempts of classifying matrix ensembles using U_N symmetry together with a finite group of involutions such as

- symmetry: $M = M^{\mathrm{T}}$,
- conjugation: $M = \bar{M}$,
- chirality and spin symmetries, that we shall not detail.

For example, the proposed classification [30, 31, 32] corresponds to the classification of symmetric spaces, and involves Dynkin diagrams. The three Wigner ensembles belong to this classification, with the GUE corresponding to the Dynkin diagram A_N. The classification of [32] contains 43 ensembles.

1.3 Universal properties

Let us give a quick glimpse of universal features of random matrix models.

1.3.1 Definition

A property is **universal** if it only depends on the matrix ensemble, and not (or almost not) on the probability measure.

The universal quantities which we will now consider are distributions of eigenvalue spacings in certain large size limits. We will start with the case of complex Hermitian matrices, with a measure built from a potential V as in (1.17). Universality then means independence from the choice of the potential V. Once it is known to be universal, a property can be studied using a quadratic potential, which leads to a Gaussian matrix integral.

Universality will arise in limits where the matrix size N is large, under specific assumptions on the large N behaviour of the potential V. We will rewrite our matrix integral as

$$\boxed{\mathcal{Z} = \int_{H_N} dM \; e^{-N \operatorname{Tr} V(M)}} \,, \tag{1.72}$$

and assume that the function $V(x)$ has a finite large N limit. Then $\mathcal{Z}$ can have a nontrivial large N limit, because the two factors dM and $e^{-N \operatorname{Tr} V(M)}$ both behave as $O(e^{N^2})$: the Lebesgue measure dM is the product of Lebesgue measures for the N^2 real components of M, the trace is a sum of N terms, and $N \operatorname{Tr} V(M) = O(N^2)$ thanks to the prefactor N. This prefactor can be absent in supermatrix models, in cases where dM effectively has fewer components due to supersymmetric cancellations.

Most known universal properties of matrix models happen to be related to integrable systems. For example, universal eigenvalue spacing distributions are typically related to tau functions of integrable systems. This is remarkable, because systems described by matrix models, such as chaotic billiards, are typically not integrable. A tentative explanation is that the statistical averaging, thanks to ergodicity, has the effect of integrating out all non-integrable degrees of freedom, and leaving only integrable degrees of freedom. This would explain why integrable systems, rather than being exceptional, are in fact commonly found in nature.

1.3.2 Macroscopic limits

Eigenvalue density and spectral curve

The **eigenvalue density** is defined as

$$\rho(x) = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} \delta(x - \lambda_i)\right] \qquad \text{(Mathematics notation)} \,, \tag{1.73}$$

$$\boxed{\rho(x) = \left\langle \frac{1}{N}\sum_{i=1}^{N} \delta(x - \lambda_i) \right\rangle} \qquad \text{(Physics notation)} \,, \tag{1.74}$$

where $\lambda_1, \ldots, \lambda_N$ are the eigenvalues of the size N matrix M. Thanks to the presence of the prefactor $\frac{1}{N}$, the eigenvalue density often has a finite limit

$$\boxed{\bar{\rho}(x) = \lim_{N\to\infty} \rho(x)} \,. \tag{1.75}$$

When it exists, this limit is called the **equilibrium density**.

The existence of the equilibrium density is not automatic. If the potential V is rational, the limit exists under some conditions on the behaviour of V near ∞ and the poles of V', such as $\lim_{x\to\infty}(V(x) - 2\log x) = \infty$ for V real. Moreover, it may happen that the limit exists only in a weak sense: instead of $\lim_{N\to\infty}\rho(x) = \bar{\rho}(x)$ for all x, we may have $\lim_{N\to\infty}\int f(x)\rho(x)dx = \int f(x)\bar{\rho}(x)dx$ for any bounded, continuous, N-independent test function $f(x)$. Most of our limits will in fact be weak limits, but we will not pursue such subtleties any further.

The equilibrium density is not a universal object. In the case of a Gaussian potential $V(M) = \frac{1}{2}M^2$, it is given by Wigner's semi-circle law (1.7). More generally, for rational potentials, the equilibrium density has a compact support, whose boundaries are called the **spectral edges**. Moreover, $y = \bar{\rho}(x)$ obeys an algebraic equation of the type

$$P(x,y) = 0 \quad \text{with} \quad P \in \mathbb{R}[x,y] \quad \text{and} \quad \deg_y P = 2\ . \tag{1.76}$$

Now, the same algebraic equation with $(x,y) \in \mathbb{C}^2$ (instead of $(x,y) \in \mathbb{R}^2$) defines the **spectral curve** of the model. The spectral curve is a Riemann surface immersed in $\mathbb{C}^2$, and can be described using algebraic geometry. The spectral curve plays a key role in random matrix theory, and more generally in integrable systems.

Two-point function

Let us then consider the **connected two-point function** in the large N limit,

$$\bar{\rho}_2(x_1,x_2) = \lim_{N\to\infty}\rho_2(x_1,x_2)\,, \tag{1.77}$$

$$\rho_2(x_1,x_2) = \left\langle \sum_{i,j}^{N}\delta(x_1-\lambda_i)\delta(x_2-\lambda_j)\right\rangle - \left\langle\sum_{i=1}^{N}\delta(x_1-\lambda_i)\right\rangle\left\langle\sum_{i=1}^{N}\delta(x_2-\lambda_i)\right\rangle\,. \tag{1.78}$$

In many cases, this two-point function has a finite $N\to\infty$ limit, which is given in terms of the eigenvalue density by a universal formula. More specifically, the two-point function only depends on the complex structure of the spectral curve, and not on its (V-dependent) embedding into $\mathbb{C}\times\mathbb{C}$. In this sense, the two-point function is a universal quantity.

Higher correlation functions

The large N **connected n-point function** is defined as

$$\boxed{\bar{\rho}_n(x_1,\dots,x_n) = \lim_{N\to\infty} N^{n-2}\left\langle\prod_{k=1}^{n}\left(\sum_{i=1}^{N}\delta(x_k-\lambda_i)\right)\right\rangle_{\mathrm{c}}}\,, \tag{1.79}$$

where the index "c" denotes the connected part or cumulant. The factor N^{n-2}, which ensures the existence of a large N limit in many (but not all) cases, can be determined by order counting in Feynman graphs. (See Chapter 2.)

The large N connected n-point function, when it exists, has a universal expression in terms of the large N one- and two-point functions. This expression is obtained by solving the topological recursion equation. (See Chapter 4.)

Expansion of the eigenvalue density

The density function (1.74) can be expanded in terms of the parameter $\frac{1}{N}$ as

$$\rho(x) = \sum_{g=0}^{\infty} N^{-2g}\bar{\rho}^{(g)}(x) + \sum_{k} N^{a_k}e^{-NA_k(x)}\,. \tag{1.80}$$

That only even powers of $\frac{1}{N}$ appear in the perturbative terms is a feature of our Hermitian ensemble: odd powers would appear in the real symmetric and Hermitian quaternionic cases. In this expansion, all perturbative and non-perturbative terms except $\bar{\rho}(x) = \bar{\rho}^{(0)}(x)$ are universal, and have universal expressions in terms of the spectral curve. For example, the functions $A_k(x)$ are periods of the spectral curve.

Similarly, the expansions of higher correlation functions are of the type

$$\rho_n(x_1, \dots, x_n) = \sum_{g=0}^{\infty} N^{-2g} \bar{\rho}_n^{(g)}(x_1, \dots, x_n) + \sum_k N^{a_k} e^{-N A_{n,k}(x_1, \dots, x_n)} \ , \tag{1.81}$$

where $\bar{\rho}_n^{(g)}$ is an algebraic function of n variables, which again has a universal expression in terms of the large N one- and two-point functions.

Expansion of the partition function

Similarly, one may consider the large N asymptotic expansion of the partition function $\mathcal{Z}$. Actually, it is not the partition function itself, but rather the **free energy** $F = \log \mathcal{Z}$, which in many cases has a nice expansion of the form

$$F = N^2 F_0 + N^0 F_1 + N^{-2} F_2 + \sum_{g=3}^{\infty} N^{2-2g} F_g + \sum_k C_k N^{a_k} e^{-N f_k} \tag{1.82}$$

The leading term F_0 is not universal, and depends on the potential $V(x)$. However, all the higher order terms are universally expressed in terms of the geometry of the spectral curve. For example F_1 is proportional to the logarithm of the Laplacian determinant on the spectral curve.

1.3.3 Microscopic limits

Bulk microscopic limit

For a point x with a non-vanishing equilibrium density $\bar{\rho}(x) \neq 0$, we define the bulk microscopic limit of the two-point function as

$$\bar{\rho}_2^{\text{micro, bulk}}(x, s) = \lim_{N \to \infty} \rho_2 \left(x, x + \frac{s}{N} \right) \ . \tag{1.83}$$

This has a universal expression in terms of the equilibrium density,

$$\bar{\rho}_2^{\text{micro, bulk}}(x, s) = - \left(\frac{\sin \pi s \bar{\rho}(x)}{\pi s \bar{\rho}(x)} \right)^2 \ . \tag{1.84}$$

This expression is related to the **sine kernel**,

$$K(x_1, x_2) = \frac{\sin(x_1 - x_2)}{x_1 - x_2} \ , \tag{1.85}$$

from which the eigenvalue spacing distribution $P(s)$ can be computed via a **Fredholm determinant**,

$$P(s) \sim - \frac{\partial^2}{\partial s^2} \det(\mathrm{Id} - \hat{K}_s) \ , \quad \hat{K}_s(f)(x) = \int_{-s}^{s} K(x, x') f(x') dx' \ . \tag{1.86}$$

This provides the exact expression for $P(s)$, of which the Wigner surmise gave an approximation. Now the sine kernel is an integrable kernel, in the sense that it can be written as

$$K(x_1, x_2) = \frac{\sum_{i,j} A_{i,j} f_i(x_1) \tilde{f}_j(x_2)}{x_1 - x_2} \ , \tag{1.87}$$

for some finite size matrix A and some functions f_i and $\tilde{f}_j$. And Fredholm determinants of integrable kernels are tau functions of integrable systems. In particular, the Fredholm determinant of the sine kernel is the tau function of the Painlevé V integrable system. (See Chapter 5.)

Edge microscopic limits

Let us consider limits of correlation functions whose arguments x_i are close to a spectral edge a. The definition of an edge microscopic limit will depend on the behaviour of the equilibrium density $\bar{\rho}(x)$, which is characterized by the algebraic equation (1.76). For a generic potential V, we have **regular edges**, such that $\bar{\rho}(x) \underset{x \to a}{\propto} \sqrt{x-a}$. For special choices of the potential, we can have $\bar{\rho}(x) \underset{x \to a}{\propto} (x-a)^{\frac{p}{2}}$ for some positive integer p. This is still a special case of the more general **critical edges**,

$$\bar{\rho}(x) \underset{x \to a}{\propto} (x-a)^{\frac{p}{q}} \ , \tag{1.88}$$

where p and q are positive integers:

Critical edges appear in multi-matrix models, which allow $\deg_y P = q \geq 2$ instead of $\deg_y P = 2$ in the algebraic equation (1.76). There are even cases when the exponent ν such that $\bar{\rho}(x) \underset{x \to a}{\propto} (x-a)^{\nu}$ is non-rational. For instance, in the $O(n)$ matrix model, we have $n = -2\cos \pi\nu$.

Near a critical edge, correlation functions with arguments x_i have finite, non-trivial limits provided

$$x_i - a \underset{N \to \infty}{\propto} N^{-\frac{q}{p+q}} \xi_i \quad , \quad \xi_i \text{ fixed} \ . \tag{1.89}$$

In the case of a regular edge $\frac{p}{q} = \frac{1}{2}$, we have $x_i - a \underset{N \to \infty}{\propto} N^{-\frac{2}{3}} \xi_i$, and the edge microscopic limit of the connected correlation function is

$$\bar{\rho}_2^{\text{micro, edge}}(\xi_1, \xi_2) \sim -K_{\text{Airy}}(\xi_1, \xi_2)^2 \ , \tag{1.90}$$

where the **Airy kernel** is written in terms of the Airy function $\mathrm{Ai}(x)$ as

$$K_{\text{Airy}}(\xi_1, \xi_2) = \frac{\mathrm{Ai}'(\xi_1)\,\mathrm{Ai}(\xi_2) - \mathrm{Ai}(\xi_1)\,\mathrm{Ai}'(\xi_2)}{\xi_1 - \xi_2} \ . \tag{1.91}$$

The Airy kernel plays an important role in deriving the Tracy–Widom law for the largest eigenvalue distribution, via the related Fredholm determinant. (See Section 5.1.3.) As a

result, the Tracy–Widom law can be expressed in terms of the tau function of the Painlevé II integrable system.

For a more general critical edge, the Airy function is replaced with a Baker–Akhiezer function $\psi_{p,q}(x)$. That function obeys a differential equation of order q whose coefficients are polynomials of degree $\leq p$, which generalizes the Airy differential equation

$$\mathrm{Ai}''(x) = x\,\mathrm{Ai}(x)\ . \tag{1.92}$$

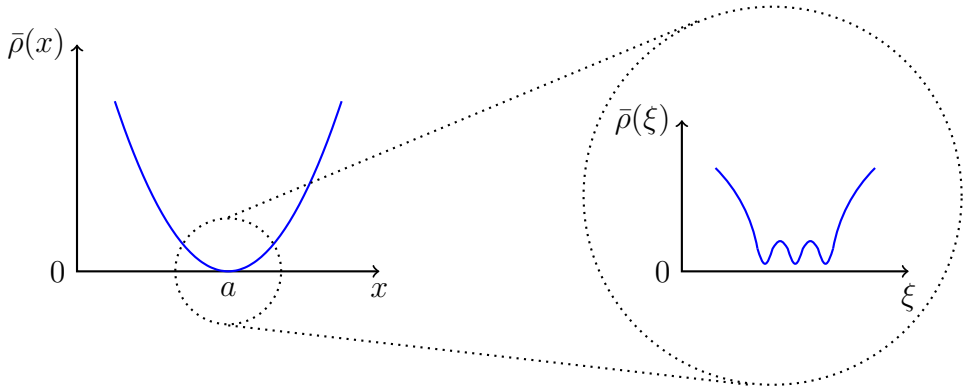
The Baker–Akhiezer functions $\psi_{p,q}(x)$ are associated with the KP ($q \geq 3$) and KdV ($q = 2$) integrable hierarchies. These integrable hierarchies are related to the (p, q) minimal models of conformal field theory, whose central charges are

$$c = 1 - 6\frac{(p-q)^2}{pq}\ . \tag{1.93}$$

In particular, the Airy case $(p, q) = (1, 2)$ corresponds to $c = -2$, whereas the pure gravity or Painlevé I case $(p, q) = (3, 2)$ corresponds to $c = 0$, and the Ising case $(p, q) = (4, 3)$ to $c = \frac{1}{2}$.

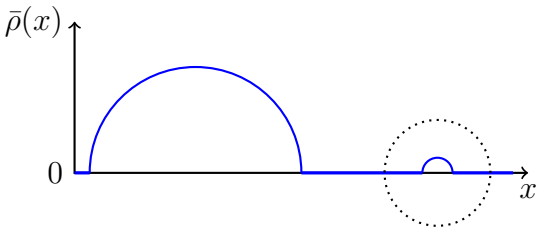
Merging and birth of edges

The microscopic limit can be used for studying the merging of edges, by focusing on the point a where two edges merge:



For generic choices of the potential V, the equilibrium density vanishes quadratically at this point, $\bar\rho(x) \propto (x-a)^2$. So the microscopic limit of the two-point function is the $(p, q) = (2, 1)$ kernel, which is the Fourier transform of the already encountered $(p, q) = (1, 2)$ Painlevé II kernel. This is a manifestation of the $(p, q) \leftrightarrow (q, p)$ duality, which is apparent in the formula (1.93) for the central charge of the corresponding conformal field theory. For special choices of the potential V, we can have $\bar\rho(x) \propto (x-a)^{2m}$, which leads to the $(p, q) = (2m, 1)$ kernel.

Similarly, when we vary V, we can observe the birth of a new edge far from the already existing edges. This gives rise to another universal microscopic regime, called the birth of a cut. Generically, the equilibrium density behaves as a square root near the newborn edge, and this regime is governed by the Hermite kernel. Other possible behaviours of the equilibrium density give rise to other integrable kernels [33, 34].



1.3.4 Non-unitary Gaussian ensembles, and β-matrix models

The universal properties of the Gaussian Unitary Ensemble ($\beta = 2$) can be generalized to the other Gaussian ensembles ($\beta \in \{1, 4\}$), where some extra complications appear. In the case of polynomial potentials, the equilibrium density $\bar{\rho}(x)$ is still an algebraic function. Higher order corrections however form an expansion in $\frac{1}{N}$ instead of $\frac{1}{N^2}$. They still have universal expressions in terms of the spectral curve.

The bulk and edge microscopic behaviours are still universal. However, in contrast to the sine or Airy kernels, the corresponding kernels are no longer of the integrable type, at least not in the classical sense of integrability. Rather, these kernels are related to quantum integrable systems, which are far less known at the present time.

Nekrasov variables and the duality $\beta \leftrightarrow \frac{4}{\beta}$

Instead of the variables N and β, let us introduce the Nekrasov variables

$$\epsilon_1 = \frac{1}{N} \qquad , \qquad \epsilon_2 = -\frac{2}{N\beta} \ . \tag{1.94}$$

These variables originate from supersymmetric gauge theory, and their identification with our matrix model variables is suggested by the AGT relation between gauge theory and two-dimensional conformal field theory, together with the relation between certain matrix integrals and the Dotsenko–Fateev integrals of conformal field theory. Now, in gauge theory, the variables ϵ_1, ϵ_2 can take arbitrary complex values. This suggests that N and β can be continued beyond their integer values.

Moreover, the gauge theory is invariant under the duality $\epsilon_1 \leftrightarrow \epsilon_2$. In conformal field theory, this corresponds to the invariance of the central charge

$$c = 1 - 6\left(\sqrt{\frac{\beta}{2}} - \sqrt{\frac{2}{\beta}}\right)^2 , \tag{1.95}$$

under $\beta \leftrightarrow \frac{4}{\beta}$. However, in matrix models, the meaning of this duality is less clear, as it corresponds to

$$\boxed{(\beta, N) \ \leftrightarrow \ \left(\frac{4}{\beta}, -\frac{N\beta}{2}\right)}. \tag{1.96}$$

Nevertheless, there is a natural matrix model quantity which is invariant under this duality: the number of independent real variables in the Gaussian ensembles E_N^β,

$$N + \frac{N(N-1)}{2}\beta = \frac{\epsilon_1 + \epsilon_2 - 1}{\epsilon_1 \epsilon_2} , \tag{1.97}$$

where N variables come from the diagonal, and $\frac{N(N-1)}{2}\beta$ variables from off-diagonal matrix elements.

The matrix model partition function is conjectured to be invariant under the duality, while observables such as the eigenvalue density or higher correlation functions are conjectured to be covariant. We will find some support for this conjecture in Section 4.1.5. In the cases $\beta = 1, 4$, the duality amounts to $S_{-2N} \leftrightarrow Q_N$. In the case $\beta = 2$, the duality amounts to $H_N \leftrightarrow H_{-N}$, which would explain why large N expansions of correlation functions involve only even powers of N.

Chapter 2

Feynman graphs

Like functional integrals, matrix integrals can be computed as perturbative series, whose terms are correlation functions in a free theory. That free theory is a Gaussian matrix model, and its correlation functions are computed in terms of Feynman graphs. Interpreting these graphs as geometrical objects gives rise to applications of matrix models to quantum gravity, string theory, statistical physics on random lattices, and combinatorics of maps.

2.1 General properties

2.1.1 Ribbon graphs

Let us consider the Hermitian Gaussian matrix model whose partition function is

$$\mathcal{Z} = \int_{H_N} dM e^{-\frac{N}{2}\operatorname{Tr} M^2} = 2^{\frac{N}{2}} \left(\frac{\pi}{N}\right)^{\frac{N^2}{2}} . \tag{2.1}$$

We are interested in correlation functions such as

$$\left\langle \operatorname{Tr} M^4 \right\rangle = \frac{1}{\mathcal{Z}} \int_{H_N} dM e^{-\frac{N}{2}\operatorname{Tr} M^2} \operatorname{Tr} M^4 . \tag{2.2}$$

Because this is just a Gaussian integral, we can apply **Wick's theorem**:

$$\left\langle MMMM \right\rangle = \overbrace{M}\overbrace{M}\overbrace{M}\overbrace{M} + \overbrace{M\,M\,M\,M} + \overbrace{M\,M\,M\,M} , \tag{2.3}$$

where each contraction of two matrices is given by the propagator

$$\overbrace{M_{ij}M_{kl}} = \left\langle M_{ij} M_{kl} \right\rangle = \frac{1}{N}\delta_{il}\delta_{jk} , \tag{2.4}$$

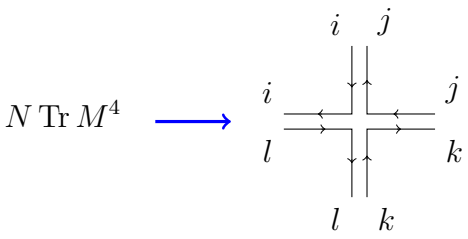
which is the inverse of the quadratic form $N \operatorname{Tr} M^2$. Thus we obtain

$$\left\langle N \operatorname{Tr} M^4 \right\rangle = \sum_{i,j,k,l} N \left\langle M_{ij} M_{jk} M_{kl} M_{li} \right\rangle , \tag{2.5}$$

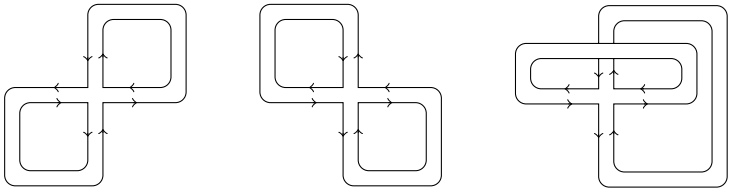
$$= \frac{1}{N} \sum_{i,j,k,l} \left(\delta_{ik}\delta_{jj}\delta_{ki}\delta_{ll} + \delta_{ii}\delta_{jl}\delta_{jl}\delta_{kk} + \delta_{il}\delta_{jk}\delta_{ji}\delta_{kl}\right) , \tag{2.6}$$

$$= N^2 + N^2 + 1 . \tag{2.7}$$

This computation can be understood in terms of Feynman graphs. Our operator $N \operatorname{Tr} M^4$ corresponds to a four-leg vertex,



Using Wick's theorem amounts to joining the four legs with one another in all possible ways. Since the propagator is trivial, joining legs forms flat ribbons:



$O(N^2)$ $O(N^2)$ $O(1)$

In a ribbon graph, the power of N is given (up to a constant) by the number of single lines. This also coincides with the number of connected components of the graph's complement when drawn on a surface. It turns out that the dominant graphs $O(N^2)$ are the **planar graphs**, which can be drawn on the plane without crossing themselves. We also have a subleading graph $O(1)$, which can be drawn on a torus but not on the plane. This can be generalized to arbitrary connected n-point correlation functions of traces of powers of M:

Theorem. ('t Hooft [14], Brézin–Itzykson–Parisi–Zuber [15])

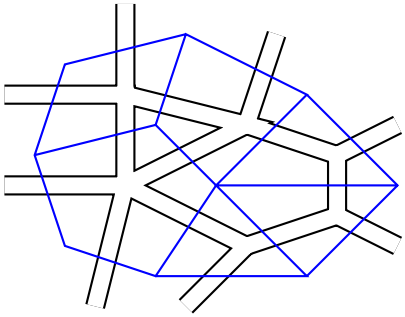
$$\boxed{\left\langle \prod_{k=1}^{n} N \frac{\operatorname{Tr} M^{p_k}}{p_k} \right\rangle_{\mathrm{c}} = \sum_{\substack{\text{connected graphs } G \\ \text{with } n \text{ vertices of valence } p_k}} \frac{N^{\chi(G)}}{\#\operatorname{Aut}(G)}}, \tag{2.8}$$

where $\#\operatorname{Aut}(G)$ is the order of the symmetry group of G, and $\chi(G) \leq 2$ is the **Euler characteristic** of G,

$$\chi(G) = \#\text{vertices} - \#\text{propagators} + \#\text{single lines}\,, \tag{2.9}$$

with $\chi(G) = 2$ for planar graphs.

Since the dual of a vertex with p_k legs is a polygon with p_k sides, the dual graph of a ribbon graph is a surface made of polygons:



The theorem can be reformulated in terms of polygonal surfaces,

$$\boxed{\left\langle \prod_{k=1}^{n} N \frac{\operatorname{Tr} M^{p_k}}{p_k} \right\rangle_{\mathrm{c}} = \sum_{\substack{\text{connected surfaces } \Sigma \\ \text{with } n \text{ polygons of size } p_k}} \frac{N^{\chi(\Sigma)}}{\#\mathrm{Aut}(\Sigma)}} \,, \tag{2.10}$$

where the Euler characteristic of a polygonal surface coincides with the characteristic of the dual graph, $\chi(\Sigma) = \chi(G)$.

Since the Euler characteristic is a topological invariant, and appears as the power of N, our formulas for correlation functions as sums of powers of N are **topological expansions**.

2.1.2 Formal matrix integrals

Given the potential

$$V(M) = \frac{1}{2} M^2 - \sum_{k=3}^{\infty} \frac{t_k}{k} M^k \,, \tag{2.11}$$

let us define a **formal matrix integral** by treating the non-Gaussian terms as perturbations,

$$\frac{1}{\mathcal{Z}} \int_{\text{formal}} dM \, e^{-N \operatorname{Tr} V(M)} \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} \frac{1}{k!} \sum_{k_1,\ldots,k_n} \left\langle \prod_{m=1}^{n} N \frac{\operatorname{Tr} M^{k_m}}{k_m} t_{k_m} \right\rangle \,. \tag{2.12}$$

This formal matrix integral is a series in $\mathbb{Q}[[t_j]]$, whose coefficients are correlation functions that can be computed using the formula (2.8),

$$\boxed{\frac{1}{\mathcal{Z}} \int_{\text{formal}} dM \, e^{-N \operatorname{Tr} V(M)} = \sum_{G} \frac{N^{\chi(G)}}{\#\mathrm{Aut}(G)} \prod_{j} t_k^{\#k-\text{vertices}}} \,. \tag{2.13}$$

The sum now includes all graphs, connected and disconnected. By standard graph combinatorics arguments, the sum over connected graphs would yield the logarithm of the formal matrix integral. Formal matrix integrals provide efficient methods for computing generating series of graphs, and algebraically encoding the combinatorial relationships among graphs. For instance, Tutte's recursion relations on the number of edges in the graphs correspond to loop equations in the formal matrix models. Deriving loop equations by integrating by parts in the formal integral is easier than finding bijections among sets of graphs. (See Chapter 4.)

The formal matrix integral in general differs from the convergent matrix integral

$$\frac{1}{\mathcal{Z}} \int_{E} dM \, e^{-N \operatorname{Tr} V(M)} \,, \tag{2.14}$$

where the ensemble E and the values of the coefficients t_k are such that the integral absolutely converges. The difference is due neither to non-perturbative terms, nor to symmetry breaking as was initially thought [35], but to the fact that integration does not always commute with the Taylor expansion of the non-Gaussian factor $e^{N \sum_{k=3}^{\infty} \frac{t_k}{k} \operatorname{Tr} M^k}$ in powers of t_k. The formal series (2.13) is actually divergent. It can sometimes be Borel resummed in some domain of the t_ks, and analytically continued to a larger domain. On a

domain where both exist, the Borel-resummed formal integral and the convergent integral may differ or agree, depending on the domain and on the ensemble E.

In general, a convergent matrix integral does not agree with the corresponding formal matrix integral, and its large N behaviour is not directly described by a topological expansion, with its integer powers of N. Nevertheless, the enumeration of ribbon graphs and the associated topological expansions are useful for studying large N asymptotics of convergent matrix integrals, as explained in Section 4.2.2. Initially the converse was expected: that N asymptotic expansions of random matrix integrals would be useful for enumerating ribbon graphs. But graph enumeration and formal series manipulations are purely algebraic, and therefore easier than large N asymptotic analysis.

2.2 Examples

We have discussed the combinatorial expansion of the Hermitian random matrix model. Let us now consider other matrix ensembles. We refer to [36] for a short review.

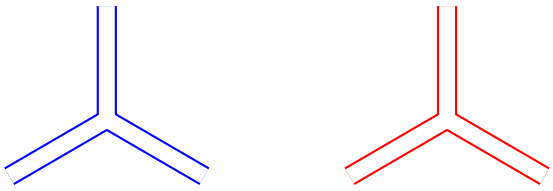
2.2.1 Multi-matrix models and colored graphs

The Ising model

Let us consider the formal two-matrix integral

$$\mathcal{Z} = \int_{\text{formal } H_N \times H_N} dM_1 dM_2 \, e^{-\frac{N}{2}\operatorname{Tr}\left(M_1^2+M_2^2-cM_1M_2\right)} \, e^{\frac{g_1}{3}\operatorname{Tr} M_1^3} \, e^{\frac{g_2}{3}\operatorname{Tr} M_2^3} \, . \tag{2.15}$$

This involves two interaction terms $\operatorname{Tr} M_1^3$ and $\operatorname{Tr} M_2^3$. In a graph, the two corresponding types of vertices can be distinguished by their colors:



$\operatorname{Tr} M_1^3$ $\qquad$ $\operatorname{Tr} M_2^3$

The color of a vertex can be interpreted as the state of a two-valued spin at this vertex. The two-matrix model therefore describes the Ising model on a random surface [37, 38].

Writing the quadratic terms in the action as

$$M_1^2 + M_2^2 - cM_1M_2 = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}^T \begin{pmatrix} 1 & -c \\ -c & 1 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} , \tag{2.16}$$

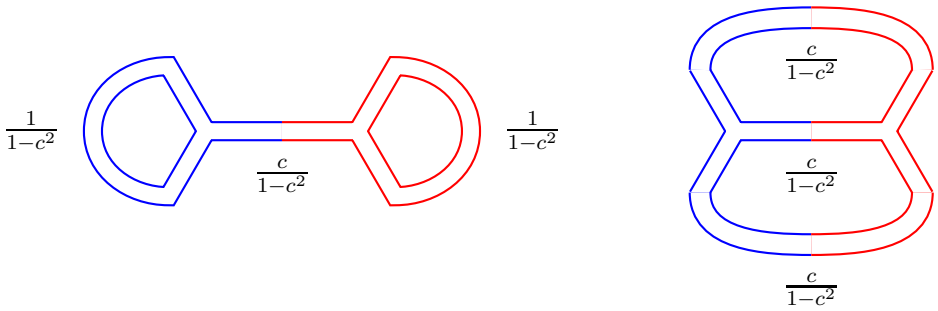
we obtain the propagator

$$\begin{pmatrix} 1 & -c \\ -c & 1 \end{pmatrix}^{-1} = \frac{1}{1-c^2} \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix} , \tag{2.17}$$

and therefore the two-point function

$$\frac{1}{N}\Big\langle (M_\alpha)_{ij} (M_\beta)_{kl} \Big\rangle = \frac{1}{1-c^2} \, \delta_{il}\delta_{jk} \times \begin{cases} 1 & (\alpha = \beta) \\ c & (\alpha \neq \beta) \end{cases} . \tag{2.18}$$

For example, the two planar graphs with two vertices of different colors are



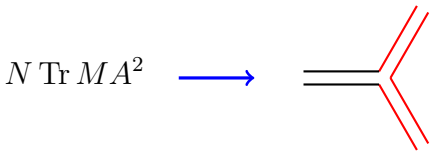
The contributions of these two graphs to the formal matrix integral differ due to the color dependence of the propagators. In a one-matrix model, the contributions of these two graphs would be equal.

The $O(n)$ matrix model

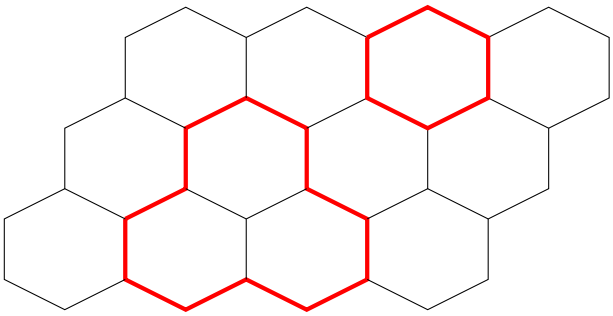
Consider the following two-matrix model,

$$\mathcal{Z} = \int_{\text{formal } H_N \times H_N} dM dA \, e^{-\frac{N}{2} \operatorname{Tr} M^2} \, e^{N \sum_{k \geq 3} \frac{t_k}{k} \operatorname{Tr} M^k} \, e^{-\frac{N}{2} \operatorname{Tr} A^2} \, e^{N \operatorname{Tr} M A^2} \, . \tag{2.19}$$

We will now color the links according to which quadratic term is used to build the propagator: $\operatorname{Tr} M^2$ or $\operatorname{Tr} A^2$. The specific feature of this model is the three-leg vertex with one M-link and two A-links,



Since the action is quadratic in A, red links must form non-intersecting loops on the resulting graphs:



This model therefore describes a loop gas. This can be generalized to the $O(n)$ model on a random surface [39, 40, 41], by having n different species for the matrix A,

$$e^{-\frac{N}{2} \operatorname{Tr} A^2} \, e^{N \operatorname{Tr} M A^2} \longrightarrow \prod_{i=1}^{n} e^{-\frac{N}{2} \operatorname{Tr} A_i^2} \, e^{N \operatorname{Tr} M A_i^2} \, . \tag{2.20}$$

It is actually possible to continue n to arbitrary non-integer values, by performing the Gaussian integrals over A_i,

$$\prod_{i=1}^{n} \int dA_i \; e^{-\frac{N}{2} \operatorname{Tr} A_i^2} \, e^{N \operatorname{Tr} M A_i^2} \propto \det(\mathrm{Id} \otimes \mathrm{Id} - M \otimes \mathrm{Id} - \mathrm{Id} \otimes M)^{-n} \, . \tag{2.21}$$

This is in particular useful for studying the replica limit $n \to 0$ and the Kosterlitz–Thouless transition $n \to 2$.

The four color problem

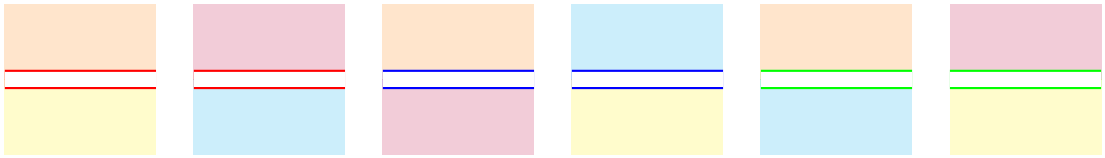
Finally, let us consider the following three-matrix integral [42, 43],

$$\int_{\text{formal } H_N^3} dA\, dB\, dC\; e^{-\frac{N}{2}\operatorname{Tr}\left(A^2+B^2+C^2\right)}\; e^{\frac{N}{3}\operatorname{Tr}(ABC+\operatorname{Tr} ACB)}\;. \tag{2.22}$$

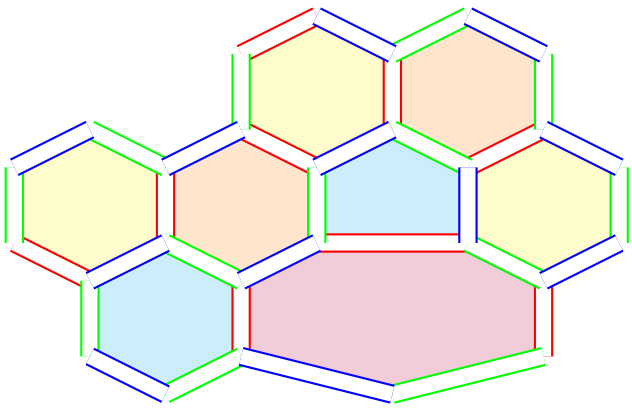
In the resulting graphs, we assign a different color to each of the propagators $\operatorname{Tr} A^2$, $\operatorname{Tr} B^2$ and $\operatorname{Tr} C^2$. Then each vertex is trivalent, and must involve all three colors:

$\operatorname{Tr} ABC$ $\qquad$ $\operatorname{Tr} ACB$

The resultant tricolor graphs can actually be mapped to the **four color problem**. The idea is that each face should have one of four colors. The colors of two adjacent faces must differ, and determine the color of the edge which separates them by a two-to-one mapping:



Starting from a tricolor graph, the colors of all faces are determined once the color of a face is given:



2.2.2 Non-Hermitian matrix model and oriented edges

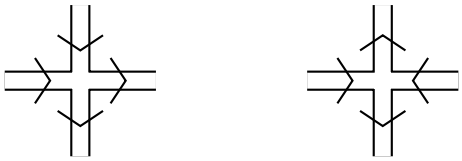
Let us consider the following formal integral over complex random matrices,

$$\mathcal{Z} = \int_{\text{formal } M_N(\mathbb{C})} dM\, e^{-N\operatorname{Tr} MM^\dagger}\, e^{t\operatorname{Tr} M^2(M^\dagger)^2}\, e^{\tilde{t}\operatorname{Tr} MM^\dagger MM^\dagger}\;. \tag{2.23}$$

We consider non-Hermitian matrices with $M \neq M^\dagger$, and this can be captured graphically by orienting the propagators:

$$\left\langle M_{ij} M^\dagger_{kl} \right\rangle = \frac{1}{N}\,\delta_{il}\delta_{jk} \qquad \longrightarrow \qquad M_{ij} \;\; \begin{matrix} i \\ j \end{matrix} \!\!\rightarrow\!\!\!\rightarrow \begin{matrix} l \\ k \end{matrix} \;\; M^\dagger_{kl}$$

The two types of vertices are therefore



$\mathrm{Tr}\, M^2 (M^\dagger)^2$ $\qquad$ $\mathrm{Tr}\, M M^\dagger M M^\dagger$

Putting these vertices on a square lattice and rotating them in all possible ways, we actually obtain six different vertices. Our matrix model therefore describes a six-vertex model on a random surface, whose continuum limit is described by a $c = 1$ conformal field theory coupled to two-dimensional gravity [44, 45].

2.2.3 Real or quaternionic matrices and twisted ribbons

Let us consider the formal integral over real symmetric matrices,

$$\mathcal{Z} = \int_{\text{formal } S_N} dM\, e^{-\frac{N}{2}\mathrm{Tr}\, M^2}\, e^{\frac{N}{4}\mathrm{Tr}\, M^4}\,. \tag{2.24}$$

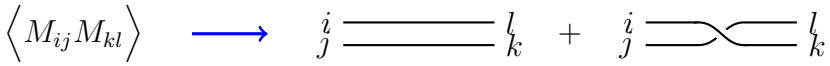
The quadratic term is

$$N\,\mathrm{Tr}\, M^2 = N \sum_{i,j=1}^{N} M_{ij} M_{ji} = N \sum_{i,j=1}^{N} M_{ij} M_{ij}\,, \tag{2.25}$$

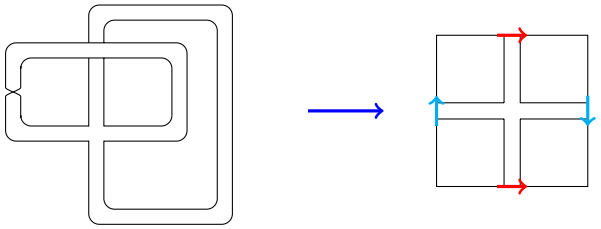
from which we deduce the propagator

$$\left\langle M_{ij} M_{kl} \right\rangle = \frac{1}{2N} \left(\delta_{il}\delta_{jk} + \delta_{ik}\delta_{jl}\right)\,. \tag{2.26}$$

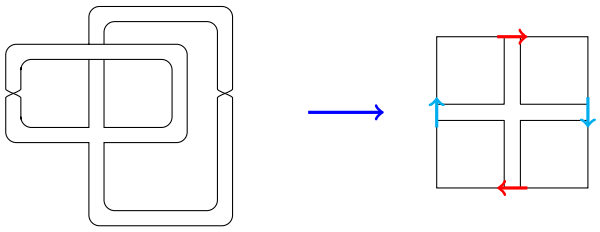
Graphically, we distinguish the two terms of the propagator by twisting (or not) the ribbon,



Graphs with twisted ribbons live on non-orientable surfaces. The surface associated to a graph is obtained by associating a square face to each four-valent vertex, and gluing squares together along edges, so that each edge is dual to a ribbon. In the case of a graph with only one vertex, we obtain the Klein bottle and the real projective plane:



Klein bottle ($\chi = 0$)



Real projective plane $\mathbb{RP}^2$
($\chi = 1$)

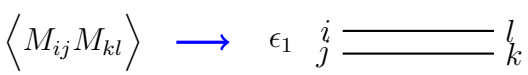
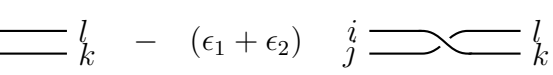
More generally, for $\beta \in \{1, 2, 4\}$, the formal integral over the Gaussian ensemble E_N^β

$$\int_{\text{formal } E_N^\beta} dM \; e^{-\frac{N\beta}{4} \operatorname{Tr} M^2} \; e^{\frac{N\beta}{2} \sum_k \frac{t_k}{k} \operatorname{Tr} M^k} \;, \tag{2.27}$$

yields the propagator

$$\left\langle M_{ij} M_{kl} \right\rangle = \epsilon_1 \delta_{il} \delta_{jk} - (\epsilon_1 + \epsilon_2) \delta_{ik} \delta_{jl} \,, \tag{2.28}$$

where we used the Nekrasov notations (1.94). Graphically,

$\left\langle M_{ij} M_{kl} \right\rangle \longrightarrow \epsilon_1$  $- (\epsilon_1 + \epsilon_2)$ 

This allows us to extend the formal integral to arbitrary values of $\beta \in \mathbb{C}$, as the generating function of weighted ribbon graphs, with weights ϵ_1 and $-(\epsilon_1+\epsilon_2)$ associated to untwisted and twisted ribbons. Twisted ribbons have a vanishing weight in the case $\beta = 2$ of Hermitian matrices.

Chapter 3

Saddle point approximation

In this Chapter we explain one method for solving matrix models in the large N limit, in other words for computing large N matrix integrals. This method is not rigorous, and it is not the most efficient method for computing higher order corrections. But it can be used for quick studies of the large N limit, where it gives an intuitive picture of the behaviour of matrix models in terms of Coulomb gases of particles. And this leads to the first hints of the deep link between random matrices and algebraic geometry.

Our main example will be the matrix integral

$$\mathcal{Z} = \int_{E_N^\beta} dM \; e^{-N\frac{\beta}{2}\operatorname{Tr} V(M)} \; , \tag{3.1}$$

where the potential V is a real polynomial of degree $d+1$, equivalently

$$d = \deg V' \; . \tag{3.2}$$

The crucial feature of this model is the invariance under conjugations of M, which allows the integral to be reduced to an integral over eigenvalues.

Our Gaussian ensembles E_N^β lead to real eigenvalues, but the results can be generalized to normal matrices with their eigenvalues in more general subsets γ of $\mathbb{C}$. Considering normal matrices is in fact necessary, because the saddle points of the integral over eigenvalues are not always real. In order to evaluate the large N limit of a matrix integral, we need γ to include the saddle points. This may require us to deform the original contour of integration over eigenvalues, whether that original contour is $\mathbb{R}$ or not. Since the integrand is analytic, such deformations do not change the integral.

3.1 The integral over eigenvalues

Since our matrix integral $\mathcal{Z}$ is invariant under conjugations by unitary matrices, critical points of $V(M)$ are hugely degenerate, and in general form a manifold of dimension $\dim U_N^\beta = O(N^2)$. The large N behaviour of $\mathcal{Z}$ is therefore not governed by these critical points. Rather, we should first reduce $\mathcal{Z}$ to an integral over eigenvalues, and then study the critical points of that integral.

3.1.1 Coulomb gas of eigenvalues

Reducing $\mathcal{Z}$ to an integral over eigenvalues (1.35) and omitting the group volume prefactor, we have

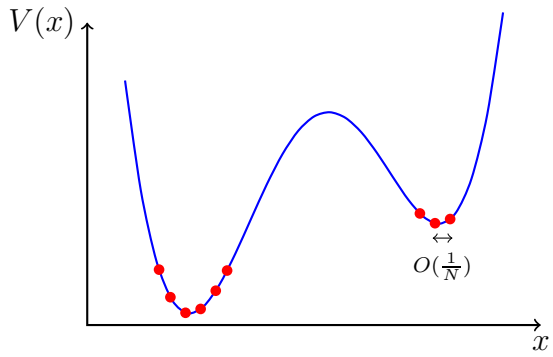
$$\mathcal{Z} \propto \left(\prod_{i=1}^N \int_\gamma d\lambda_i \right) e^{-N^2 \frac{\beta}{2} S(\lambda_1,\ldots,\lambda_N)} \; , \tag{3.3}$$

where the action is

$$\boxed{S(\lambda_1,\dots,\lambda_N) = \frac{1}{N}\sum_{i=1}^{N} V(\lambda_i) - \frac{1}{N^2}\sum_{i<j}^{N} \log(\lambda_i-\lambda_j)^2} \ . \tag{3.4}$$

This action is β-independent, and the value of β plays no role in the large N limit. Moreover, assuming V to be N-independent, a large N limit of our action has a good chance to exist. This is because we followed the prescription of Section 1.3.1 when writing our matrix integral $\mathcal{Z}$, and included a prefactor N in the exponent.

Our action describes a **Coulomb gas** of eigenvalues, which can be thought of as point particles in the one-dimensional space $\mathbb{R}$, subject to a repulsive Coulomb interaction. At equilibrium, the eigenvalues typically gather in potential wells, but cannot all gather at the minimum due to their mutual repulsion, which keeps them at distances $O(\frac{1}{N})$. So the eigenvalues tend to occupy a finite segment around each potential well:



(3.5)

The **saddle point approximation** of $\mathcal{Z}$ consists in localizing the integral at the critical points of the action. The critical points are defined by the N saddle point equations $\frac{\partial S}{\partial \lambda_i} = 0$, which are equivalent to the N coupled algebraic equations

$$V'(\lambda_i) = \frac{2}{N}\sum_{j\neq i}\frac{1}{\lambda_i-\lambda_j} \quad \text{for} \quad i=1,\dots,N \, . \tag{3.6}$$

(For an appropriate choice of the potential V, this is formally identical to the Bethe ansatz equations for the Jaynes–Cummings–Gaudin model. This is not a mere accident, and it is a hint of the deep link between random matrices and integrable systems).

Without proving it, let us mention that the number of critical points, up to permutations of the eigenvalues, is

$$\#\left\{\{\lambda_1,\dots,\lambda_N\}\middle|\frac{\partial S}{\partial\lambda_i}=0\right\}/\mathfrak{S}_N = \binom{N+d-1}{N} = d_N \ . \tag{3.7}$$

The number of critical points coincides with the number d_N of homologically independent convergent integrals with the action S (1.45), which emerged from our study of filling fractions. This suggests an analogy between the choice of a critical point, and the choice of an integration contour, for a given potential. We will explain in this Chapter how to find the solution that dominates the large N asymptotics – in other words, the equilibrium density of eigenvalues.

3.1.2 The large N limit as a WKB approximation

Given a solution $\{\lambda_i\}$ of the saddle point equations, we define the **resolvent**

$$\boxed{\omega(x) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{x-\lambda_i}} \ . \tag{3.8}$$

Using the saddle point equations, we compute

$$\omega(x)^2 + \frac{1}{N}\omega'(x) = \frac{1}{N^2}\left(\sum_{i,j=1}^{N}\frac{1}{x-\lambda_i}\frac{1}{x-\lambda_j} - \sum_{i=1}^{N}\frac{1}{(x-\lambda_i)^2}\right) , \tag{3.9}$$

$$= \frac{2}{N^2}\sum_{i=1}^{N}\frac{1}{x-\lambda_i}\sum_{j\neq i}\frac{1}{\lambda_i-\lambda_j} , \tag{3.10}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\frac{V'(\lambda_i)}{x-\lambda_i} , \tag{3.11}$$

$$= V'(x)\omega(x) - P(x) , \tag{3.12}$$

where we introduced

$$\boxed{P(x) = \frac{1}{N}\sum_{i=1}^{N}\frac{V'(x)-V'(\lambda_i)}{x-\lambda_i}} . \tag{3.13}$$

which is a polynomial of x of degree $d-1$. Actually, if the potential V is rational, then $P(x)$ is again a rational function of x, with a divisor

$$\mathrm{div}(P) \leq (\deg_\infty V' - 1)\cdot[\infty] + \sum_{\alpha\neq\infty}(\deg_\alpha V')\cdot[\alpha] , \tag{3.14}$$

and therefore again $\deg P \leq d-1$, with $d = \deg V'$.

The equation (3.12) is a closed equation for $\omega(x)$, provided we know $P(x)$. In general, we know a good deal of information on $P(x)$: that it is a polynomial (or rational function) of degree $d-1$, and that its leading coefficient is the same as that of $V'(x)$. This fully determines $P(x) = V''(x)$ if $d = 1$, and leaves $d-1$ unknown coefficients in general. For the moment we will work with arbitrary values of these coefficients, and consider P as known. Later, we will explain how to determine P.

The differential, Riccati equation (3.12) for $\omega(x)$ can be rewritten as the Schrödinger equation

$$\frac{1}{N^2}\psi'' - \frac{1}{N}V'\psi' + P\psi = 0 , \tag{3.15}$$

for the function $\psi(x)$ defined by

$$\omega = \frac{1}{N}\frac{\psi'}{\psi} . \tag{3.16}$$

Equivalently, $\psi(x)$ is the characteristic polynomial

$$\psi(x) = \prod_{i=1}^{N}(x-\lambda_i) = \det(x-M) . \tag{3.17}$$

The large N limit is now interpreted as the WKB approximation for the Schrödinger equation. In this limit, we neglect the $\frac{1}{N}\omega'$ term, and the Riccati equation (3.12) reduces to the algebraic equation

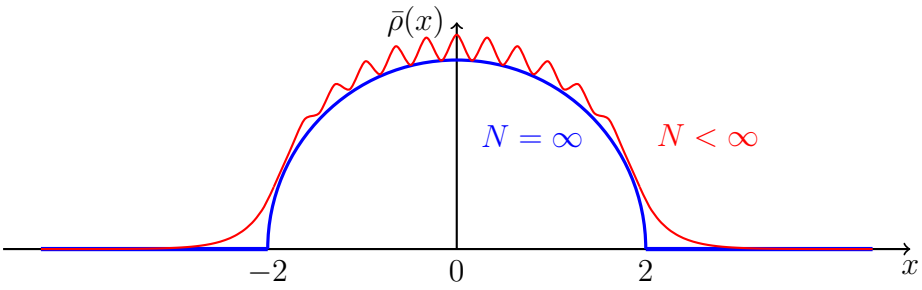
$$\boxed{\bar\omega^2 - V'\bar\omega + \bar P = 0 \quad \text{with} \quad \begin{cases} \bar\omega = \lim\limits_{N\to\infty}\omega , \\ \bar P = \lim\limits_{N\to\infty} P , \end{cases}} \tag{3.18}$$

whose solution is

$$\bar{\omega} = \frac{1}{2}\left(V' - \sqrt{V'^2 - 4\bar{P}}\right) \ . \tag{3.19}$$

As we will see in the next Subsection, the density of eigenvalues $\bar{\rho}$ can be deduced from $\bar{\omega}$.

In the case of a quadratic potential V, we have $\bar{P} = P = V''$, and the density of eigenvalues which follows from the above solution $\bar{\omega}$ is Wigner's semi-circle law. For finite values of N, the Schrödinger equation yields corrections to this law, and the density is modified as follows:



(3.20)

Corrections to the semi-circle law are oscillations of frequency $O(N)$ and amplitude $O(\frac{1}{N})$ on the semi-circle's support, and exponentially decreasing outside the support, as usual in WKB approximations.

For general potentials, $\bar{\omega}$ depends on the $d-1$ unknown coefficients of the polynomial $\bar{P}$. We could compute the action (3.4) as a function of these coefficients, and fix their values by extremizing the action. For the moment we assume that this yields unique values for these coefficients. We will come back to this point in Section 3.2.3.

3.2 Determining the equilibrium density

In this Section, we give a rather pedestrian method for computing the equilibrium density of eigenvalues $\bar{\rho}$. We will first determine $\bar{\rho}$ in terms of its support $\operatorname{supp}\bar{\rho}$, and of the polynomial $\bar{P}$. We will then determine these two pieces of data.

3.2.1 Functional saddle point equation

The action (3.4), which we defined as a function of the eigenvalues λ_i, can be rewritten in the large N limit as a functional of the equilibrium density $\bar{\rho}(x)$,

$$S[\bar{\rho}] = \int_{\mathbb{R}} dx\, \bar{\rho}(x) V(x) - \fint_{\mathbb{R}^2} dx dx'\, \bar{\rho}(x)\bar{\rho}(x') \log|x-x'| + \ell\left(1 - \int_{\mathbb{R}} dx\, \bar{\rho}(x)\right) \ , \tag{3.21}$$

where the variable ℓ is a Lagrange multiplier, which ensures that the total density mass is one, and allows us to consider S as a functional over not necessarily normalized densities. For later use we regularize the double integral by taking its principal part, although it is convergent. The generalization to a normal matrix model, whose eigenvalues belong to a union of contours $\gamma = \cup\gamma_i$ with the filling fractions ϵ_i, would be

$$S[\bar{\rho}] = \int_{\gamma} dx\, \bar{\rho}(x) V(x) - \fint_{\gamma^2} dx dx'\, \bar{\rho}(x)\bar{\rho}(x') \log|x-x'| + \sum_i \ell_i \left(\epsilon_i - \int_{\gamma_i} dx\, \bar{\rho}(x)\right) \ . \tag{3.22}$$

In terms of the Fourier-transformed density $\tilde{\bar{\rho}}(k) = \int_{\mathbb{R}} dx\, e^{ikx} \bar{\rho}(x)$, the quadratic term becomes

$$- \fint_{\mathbb{R}\times\mathbb{R}} dx dx'\, \bar{\rho}(x)\bar{\rho}(x') \log|x-x'| = \int_{\mathbb{R}} \frac{dk}{|k|}\, \tilde{\bar{\rho}}(k)\tilde{\bar{\rho}}(-k) = \frac{1}{2}\int_0^\infty \frac{dk}{k}\, |\tilde{\bar{\rho}}(k)|^2 \ , \quad (3.23)$$

where we used the condition $\bar{\rho}(x) \in \mathbb{R}$. This shows that $S[\bar{\rho}]$ is a real convex functional, which therefore has a unique minimum in the space of real positive densities.

The saddle points of $S[\bar{\rho}]$ are defined by the equation $\frac{\delta S}{\delta \bar{\rho}(x)} = 0$. Using the **effective potential**,

$$V_{\text{eff}}(x) = V(x) - 2 \fint_{\mathbb{R}} dx'\, \bar{\rho}(x') \log|x-x'| \ , \quad (3.24)$$

the saddle point equation can be written as

$$\forall\, x \in \text{ supp}\, \bar{\rho}\,, \qquad V_{\text{eff}}(x) = \ell \ . \quad (3.25)$$

Derivating with respect to x, we obtain

$$\forall\, x \in \text{ supp}\, \bar{\rho}\,, \qquad V'(x) = 2 \fint_{\mathbb{R}} \frac{dx'}{x-x'}\, \bar{\rho}(x') \ , \quad (3.26)$$

which is identical to the continuum limit of the saddle point equation (3.6) for the eigenvalues. The derivative of the saddle point equation can be written as a **Riemann–Hilbert equation**,

$$\boxed{\forall\, x \in \text{ supp}\, \bar{\rho}\,, \qquad V'(x) = \bar{\omega}(x+i0) + \bar{\omega}(x-i0)} \ , \quad (3.27)$$

for the **Stieltjes transform** $\bar{\omega}(x)$ of the measure $\bar{\rho}(x)dx$,

$$\boxed{\bar{\omega}(x) = \int_{\text{supp}\,\bar{\rho}} \frac{dx'}{x-x'}\, \bar{\rho}(x')} \ . \quad (3.28)$$

This definition of $\bar{\omega}(x)$ agrees with the earlier definition (3.18) as the large N limit of the resolvent. Then $\bar{\omega}(x)$ is a holomorphic function of x outside of the support of $\bar{\rho}$, which behaves near $x = \infty$ as

$$\bar{\omega}(x) \underset{x\to\infty}{=} \frac{1}{x} + O\left(\frac{1}{x^2}\right) \ , \quad (3.29)$$

and which has a jump on the support of $\bar{\rho}$:

$$\boxed{\bar{\omega}(x+i0) - \bar{\omega}(x-i0) = -2\pi i \bar{\rho}(x)} \ . \quad (3.30)$$

Together with eq. (3.27), this implies that the function

$$\bar{P}(x) = V'(x)\bar{\omega}(x) - \bar{\omega}(x)^2 \ , \quad (3.31)$$

is analytic on $\mathbb{C}$. This is because

$$\bar{P}(x+i0) - \bar{P}(x-i0) = -\Big(\bar{\omega}(x+i0) - \bar{\omega}(x-i0)\Big)\Big(\bar{\omega}(x+i0) + \bar{\omega}(x-i0) - V'(x)\Big) = 0 \ , \quad (3.32)$$

due to one factor vanishing on the support of $\bar{\rho}$, and the other factor vanishing elsewhere. Since by definition $\bar{P}(x) \underset{x\to\infty}{\sim} \frac{V'(x)}{x}$, it follows that $\bar{P}(x)$ is a polynomial of degree $\deg \bar{P} = d-1$, which actually coincides with the previously defined $\bar{P}(x)$ (3.18).

The polynomial $\bar{P}$ is supposed to be determined by the saddle point equation, while any choice of $\bar{P}$ gives rise to a solution $\bar{\omega}(x)$ of the Riemann–Hilbert equation. The equilibrium density is then deduced via (3.30),

$$\boxed{\bar{\rho} = \frac{1}{2\pi}\sqrt{4\bar{P} - (V')^2} \times \mathbf{1}_{\operatorname{supp}\bar{\rho}}} \,. \tag{3.33}$$

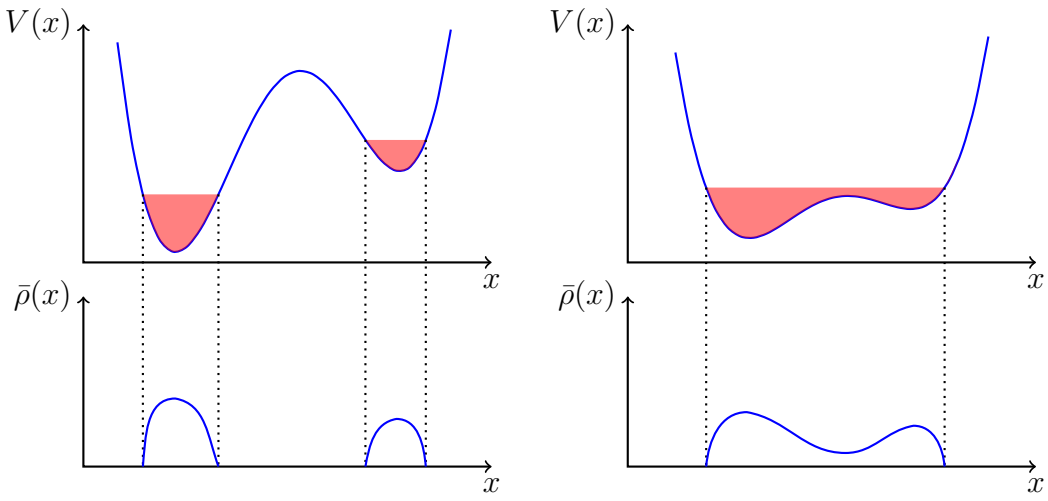
This determines the equilibrium density in terms of its support $\operatorname{supp}\bar{\rho}$, and of the polynomial $\bar{P}$.

3.2.2 Determining the support

In the case where the eigenvalues and the polynomial V are real, we have

$$\operatorname{supp}\bar{\rho} = \left\{4\bar{P} - (V')^2 > 0\right\} \subset \mathbb{R} \,, \tag{3.34}$$

so that the equilibrium density is real on its support. This support is a union of compact intervals, which correspond to the wells of the potential where the eigenvalues gather. The number of intervals is sensitive to the precise shape of the potential. For example, in the case of a quartic potential $\deg V = 4$ with two wells, we can have two intervals if the wells are deep enough, or one single interval if they are not:



Since $\deg\left(4\bar{P} - (V')^2\right) = 6$, it may seem that the support can be made of three intervals. As we will shortly show, this would however be incompatible with the positivity of the equilibrium density.

If on the other hand the potential V is not real, then $\bar{\rho}$ can be supported on a number of arcs in the complex plane, and needs not to be real. What should be real is the one-form $\bar{\rho}(x)dx$, so that its integrals on subsets of the support are real (and even actually positive). The equilibrium density $\bar{\rho}$ and its support are in principle determined by extremizing the action $S[\bar{\rho}]$ with respect to both of them, while keeping $\bar{\rho}(x)dx$ positive on $\operatorname{supp}\bar{\rho}$. In contrast to the case where V is real, the action is not a convex functional, and its saddle point is a true saddle, not a minimum. Finding the saddle point, and showing that it is unique, now requires solving a min-max extremization problem, where $S[\bar{\rho}]$ should be minimized with respect to $\bar{\rho}$, and maximized with respect to the support. Such min-max extremization problems are in general very difficult. Here we will give the solution but omit the proof, which involves algebraic geometry beyond the scope of this text [46].

So let us give the recipe for determining $\operatorname{supp}\bar{\rho}$, assuming we know the polynomial $\bar{P}$. The expressions for $\bar{\rho}$ and $\bar{\omega}$ in terms of the polynomial $\bar{P}$ involve the square root of the

polynomial $(V')^2 - 4\bar{P}$. Let us isolate the perfect square factor in this polynomial, and write it as

$$\boxed{(V')^2 - 4\bar{P} = M^2\sigma}\,, \tag{3.35}$$

so that the equations (3.33) and (3.31) can be rewritten as

$$\boxed{\bar{\rho} = -\frac{1}{2\pi} M\sqrt{-\sigma} \times \mathbf{1}_{\operatorname{supp}\bar{\rho}} \quad , \quad \bar{\omega} = \frac{1}{2}\left(V' - M\sqrt{\sigma}\right)}\,. \tag{3.36}$$

More specifically, $M(x)$ is a polynomial, while $\sigma(x)$ is a product over all roots of $(V')^2 - 4\bar{P}$ with odd multiplicities,

$$\sigma(x) = \prod_{j=1}^{2s}(x - a_i)\ . \tag{3.37}$$

The number of such roots is always even, because $\deg\left((V')^2 - 4\bar{P}\right) = \deg(V')^2$ is even.

If the potential and eigenvalues are real, and if moreover the roots a_i are real, we can assume $a_i < a_{i+1}$ and write

$$\operatorname{supp}\bar{\rho} = \bigcup_{i=1}^{s}(a_{2i-1}, a_{2i})\ . \tag{3.38}$$

The constraint $\bar{\rho} > 0$ leads to a bound on the number s of intervals. This is because $\sqrt{-\sigma}$ changes sign from one interval to the next. So $\bar{\rho} > 0$ implies that M has a zero in (a_{2i}, a_{2i+1}), so that $\deg M \geq s - 1$. This implies $s \leq \frac{d}{2}$, which is consistent with the intuitive expectation that the number of intervals cannot exceed the number of potential wells.

In more general cases, the connected components of $\operatorname{supp}\bar{\rho}$ are paths between the zeros a_i of σ, on which $\bar{\rho}(x)dx$ is a positive measure. This implies that $\operatorname{supp}\bar{\rho}$ is a subset of the **spectral network**

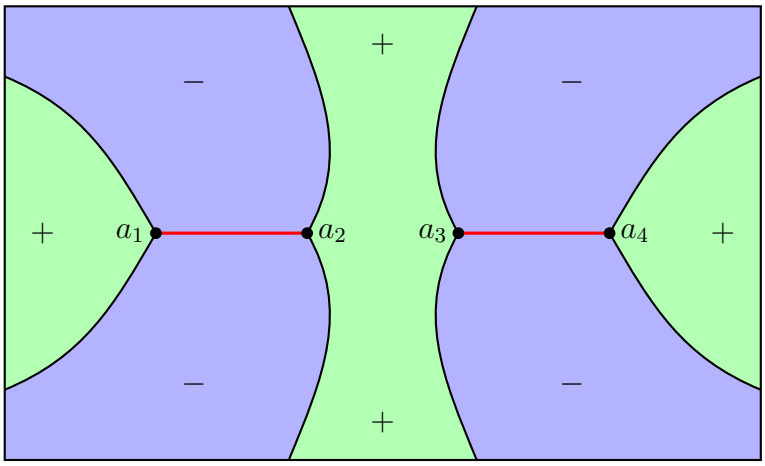
$$\boxed{\mathcal{S} = \bigcup_{i=1}^{2s}\{\Re g_i(x) = 0\} \quad \text{with} \quad g_i(x) = \int_{a_i}^{x} M(x')\sqrt{\sigma(x')}dx'}\,. \tag{3.39}$$

The spectral network is a graph in the complex plane, which is not necessarily connected. Since $g_i(x) \underset{x\to a_i}{\sim} \frac{M(a_i)\sqrt{\sigma'(a_i)}}{3\pi}(x - a_i)^{\frac{3}{2}}$, the spectral network has a trivalent vertex at a_i, except in special cases where $M(a_i) = 0$. Each one of the edges starting from a_i either ends up at a_j with $j \neq i$, or goes to infinity in a direction where $\Re V(x) \underset{x\to\infty}{\to} 0$.

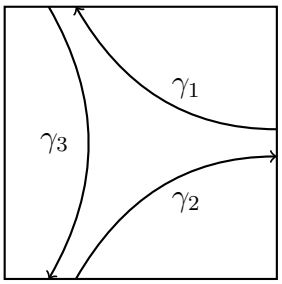
To determine $\operatorname{supp}\bar{\rho}$ as a subset of the spectral network $\mathcal{S}$, consider the complement $\mathbb{C} - \mathcal{S}$ of $\mathcal{S}$. It can be proved that the connected components of $\mathbb{C} - \mathcal{S}$ are all non-compact, and therefore reach ∞. To each connected component, let us assign the sign $\pm$ such that $\Re V(x) \underset{x\to\infty}{\to} \pm\infty$ in that domain. (That sign is always well-defined, even if the component reaches infinity in several sectors.) Then $\operatorname{supp}\bar{\rho}$ is the union of the edges of $\mathcal{S}$ that separate two negative components. The edges between a negative and a positive component are therefore not included, and it can be shown that two positive components cannot be contiguous.

It is customary to think of positive components as land, and of negative components as seas. Then $\operatorname{supp}\bar{\rho}$ is the union of all bridges – the edges of $\mathcal{S}$ that separate two seas.

In general $\operatorname{supp}\bar{\rho}$ is a disjoint union of connected trivalent trees. When the potential and eigenvalues are real, these trees reduce to segments. For example, in the case of a quartic potential $V(x) = x^4 + \cdots$, the edges go to infinity in the eight directions of the type $\frac{\pi}{8} + n\frac{\pi}{4}$. Assuming we have $2s = 4$ roots a_i, the seas, lands and bridges looks as follows:



We now give the simplest possible example of a support with a nontrivial trivalent vertex. Consider a normal matrix model with the cubic potential $V(x) = \frac{4}{3}x^3$, and with eigenvalues on a path γ. For the matrix integral to converge, γ should belong to the homology space associated to our potential. This homology space is generated by the following three paths:



Let us look for γ as a combination of these three paths that respects the symmetry of the potential under the rotation by $\frac{2\pi}{3}$. While the obvious invariant combination $\gamma_1+\gamma_2+\gamma_3 = 0$ is trivial, the nontrivial combination

$$\gamma = \sum_{k=1}^{3} e^{\frac{2\pi i}{3}k}\gamma_k \ , \tag{3.40}$$

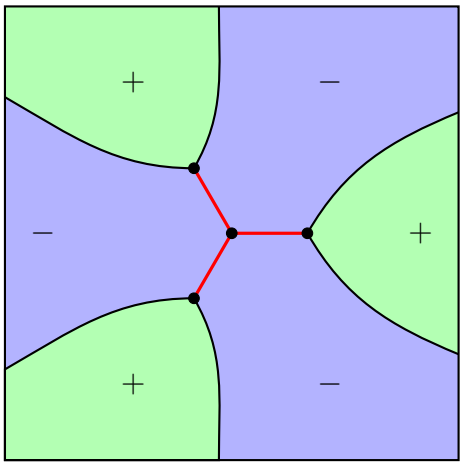
is projectively invariant. Then we expect the spectral network to be invariant as well. This requires that $(V')^2 - 4\bar{P}$ be projectively invariant. Given $V'(x) = 4x^2$, this implies that the constant coefficient of the degree one polynomial $\bar{P}(x)$ vanishes, and we have

$$V'(x)^2 - 4\bar{P}(x) = 16x(x^3-1) \quad \Rightarrow \quad \{a_i\} = \left\{0, 1, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}\right\} . \tag{3.41}$$

The spectral network $\mathcal{S}$ turns out to be connected, and therefore determined by only one function

$$g_0(x) = 4\int_0^x dx' \sqrt{x'(x'^3-1)} = \frac{1}{3}\left(\frac{1}{\zeta} - \zeta + 2\log\zeta\right) , \quad \text{with} \quad 2x^{\frac{3}{2}} = \sqrt{\zeta} + \frac{1}{\sqrt{\zeta}} . \tag{3.42}$$

Then $\mathcal{S} = \{\Re g_0(x) = 0\} = \{|\zeta| = 1\} \cup \left\{\cos(\arg\zeta) = \frac{\log|\zeta|}{|\zeta|-|\zeta|^{-1}}\right\}$. In the complex x-plane, the circle $\{|\zeta| = 1\}$ corresponds to the three segments $[0,1]$, $[0, e^{\frac{2\pi i}{3}}]$ and $[0, e^{\frac{4i\pi}{3}}]$. These three segments are bridges, whose union is $\operatorname{supp}\bar{\rho}$:



Finally, let us consider the example of the ensemble U_N of unitary matrices with the Haar measure (1.47). This can be viewed as a normal matrix ensemble with the potential $V(x) = \log x$. The number of independent integration contours is $\deg V' = 1$. The topologically unique contour is a loop around $x = 0$. We can in principle have $\bar{P}(x) = \frac{a}{x}$ for an arbitrary coefficient a, but preserving the symmetry under rotations around $x = 0$, and therefore requiring that $(V')^2 - 4\bar{P}$ be projectively invariant, implies $a = 0$. The equilibrium density (3.33) is

$$\bar{\rho}(x)dx = \frac{1}{2\pi i}\frac{dx}{x} = \frac{d\theta}{2\pi} + \frac{1}{2\pi i}\frac{dr}{r} \quad \text{with} \quad x = re^{i\theta} \ . \tag{3.43}$$

The positivity of the density implies $\frac{dr}{r} = 0$, and the contour must be a circle. Alternatively, the result that eigenvalues are uniformly distributed on a circle could have been obtained as a direct consequence of the U_N symmetry.

3.2.3 Determining the polynomial $\bar{P}(x)$

The Riemann–Hilbert equation means that the effective potential V_{eff} has a vanishing derivative on $\operatorname{supp}\bar{\rho}$, and is therefore constant on each connected component of $\operatorname{supp}\bar{\rho}$. The saddle point equation (3.25) implies in addition that the effective potential must take the same value ℓ on every connected component. (More generally, in a normal matrix model (3.22), we would have $V_{\text{eff}}(x) = \ell_i$ on all connected components associated to the path γ_i.) Now, by the definitions of the effective potential and of the resolvent $\bar{\omega}$, the derivative of the effective potential outside $\operatorname{supp}\bar{\rho}$ is

$$V'_{\text{eff}} = V' - 2\bar{\omega} = M\sqrt{\sigma} \ , \tag{3.44}$$

where we used eq. (3.36). The saddle point equation then implies

$$\forall i \in \{1, \cdots s-1\} \ , \quad \int_{a_{2i}}^{a_{2i+1}} dx\, V'_{\text{eff}}(x) = \int_{a_{2i}}^{a_{2i+1}} dx\, M(x)\sqrt{\sigma(x)} = 0 \ . \tag{3.45}$$

This yields $s-1$ equations for $\bar{P}$. Moreover, requiring that $(V')^2 - 4\bar{P}$ has $d-s$ double zeros as follows from eq. (3.35), yields another $d-s$ equations. We thus have a total of $d-1$ equations, and that matches the number of unknown coefficients of $\bar{P}$. We will admit that these equations determine $\bar{P}$, even when the roots a_i are not real.

Actually, once the support or equivalently σ is known, the resolvent $\bar{\omega}$ is then given by the **Tricomi relation**,

$$\boxed{\bar{\omega}(x) = \frac{1}{2\pi i} \int_{\operatorname{supp}\bar{\rho}} dx' \sqrt{\frac{\sigma(x)}{\sigma(x')}} \frac{V'(x')}{x - x'}} \ . \tag{3.46}$$

The proof uses the behaviour of $\bar{\omega}(x)$ near $x = \infty$, the expression (3.36) for $\bar{\omega}$ in terms of σ and a polynomial M, and the relation $\oint_{\operatorname{supp}\bar{\rho}} \frac{f(x')}{\sqrt{\sigma(x')}} = -2 \int_{\operatorname{supp}\bar{\rho}} \frac{f(x')}{\sqrt{\sigma(x')}}$ for any function f that is analytic near $\operatorname{supp}\bar{\rho}$:

$$\frac{\bar{\omega}(x)}{\sqrt{\sigma(x)}} = \frac{1}{2\pi i} \oint_x \frac{dx'}{x' - x} \frac{\bar{\omega}(x')}{\sqrt{\sigma(x')}} = -\frac{1}{2\pi i} \oint_{\operatorname{supp}\bar{\rho}} \frac{dx'}{x' - x} \frac{\bar{\omega}(x')}{\sqrt{\sigma(x')}} \ , \tag{3.47}$$

$$= -\frac{1}{4\pi i} \oint_{\operatorname{supp}\bar{\rho}} \frac{dx'}{x' - x} \left[\frac{V'(x')}{\sqrt{\sigma(x')}} - M(x') \right] = -\frac{1}{4\pi i} \oint_{\operatorname{supp}\bar{\rho}} \frac{dx'}{x' - x} \frac{V'(x')}{\sqrt{\sigma(x')}} \ , \tag{3.48}$$

$$= \frac{1}{2\pi i} \int_{\operatorname{supp}\bar{\rho}} \frac{dx'}{x' - x} \frac{V'(x')}{\sqrt{\sigma(x')}} \ . \tag{3.49}$$

In the case where $\operatorname{supp}\bar{\rho} = (a, b)$ reduces to one interval, we can actually determine its boundaries a and b by analyzing the large x behaviour of the Tricomi relation,

$$\bar{\omega}(x) \underset{x\to\infty}{=} \frac{1 - \frac{a+b}{2x}}{2\pi i} \int_a^b dx' \frac{1 + \frac{x'}{x}}{\sqrt{(x' - a)(x' - b)}} V'(x') + O\left(\frac{1}{x^2}\right) \ . \tag{3.50}$$

Comparing with the known behaviour of $\bar{\omega}(x)$ (3.29), we obtain two equations for the two unknowns a and b,

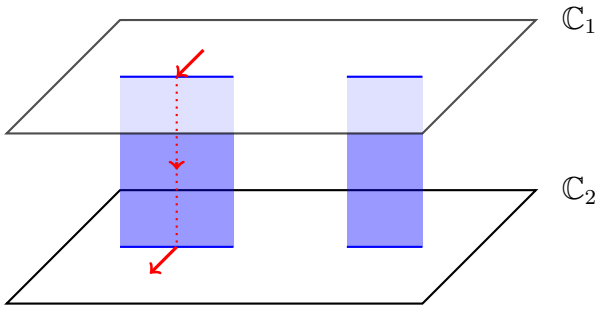
$$\int_a^b dx' \frac{V'(x')}{\sqrt{(x' - a)(x' - b)}} = 0 \quad , \quad \frac{1}{2\pi i} \int_a^b dx' \frac{x' V'(x')}{\sqrt{(x' - a)(x' - b)}} = 1 \ . \tag{3.51}$$

3.3 Algebraic geometry

3.3.1 The spectral curve

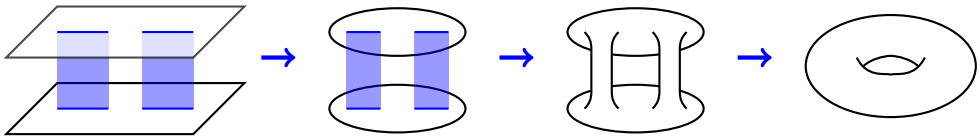
Definition

The resolvent $\bar{\omega}(x)$ is defined on the complex plane, minus the support of $\bar{\rho}$, where it has a discontinuity. Describing the behaviour of such functions was Riemann's original motivation for introducing Riemann surfaces. The idea is that $\bar{\omega}(x)$ is defined on a two-sheeted cover of our original complex plane – the two sheets correspond to the two solutions of the quadratic equation (3.18). The support of $\bar{\rho}$ is interpreted as a collection of cuts on both sheets, and it is possible to smoothly move across a cut provided one changes sheets:



In this picture we illustrated the case of the double well potential (3.5). In this case the eigenvalues gather in the two wells, which in the large N limit give rise to two cuts.

What is the intrinsic geometry of the resulting manifold? Our two complex planes are actually **Riemann spheres** $\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, as nothing special happens at $x = \infty$. The cuts which connect them can be considered as smooth tubes. If we have one cut, then the resulting Riemann surface is again a Riemann sphere. If we have two cuts, then the resulting Riemann surface is a torus:



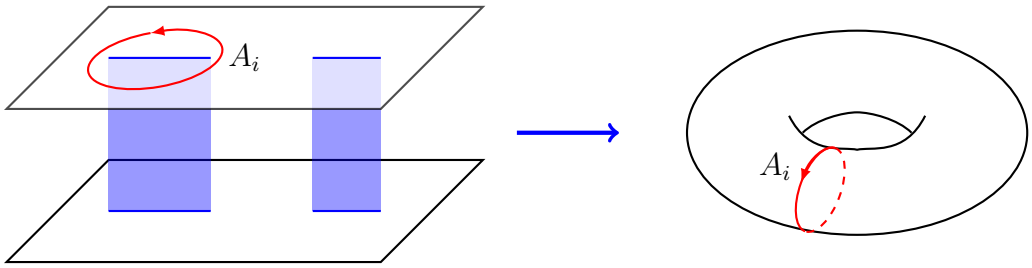
More generally, if we have s cuts, the resulting Riemann surface has the genus $g = s - 1$. That surface is called the **spectral curve** of our model – the manifold where the resolvent lives as a smooth function. This manifold can alternatively be described as the subset of $\bar{\mathbb{C}} \times \bar{\mathbb{C}}$ defined by the algebraic equation

$$\boxed{y^2 - V'(x)y + \bar{P}(x) = 0}\,. \tag{3.52}$$

This provides a precise definition of the spectral curve, which we introduced in Section 1.3.2. The topology and complex structure of the spectral curve are universal properties, whereas its embedding in $\bar{\mathbb{C}} \times \bar{\mathbb{C}}$ is not universal.

Periods

Let A_i be a smooth closed non-contractible cycle of the Riemann surface, surrounding the cut (a_{2i-1}, a_{2i}) in the first sheet:

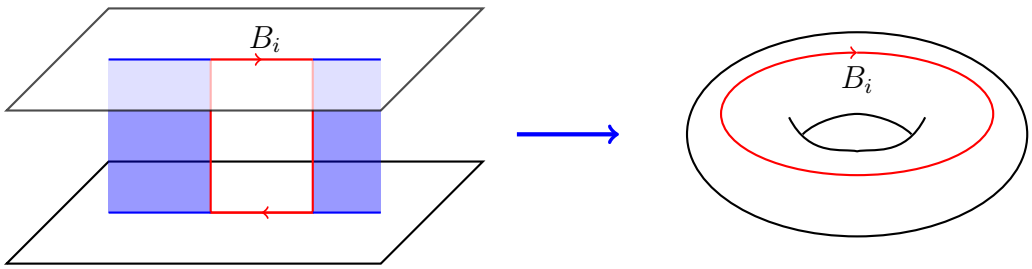


The integral of the equilibrium density on (a_{2i-1}, a_{2i}), which we call a **filling fraction**, can then be rewritten as the integral of the resolvent over the cycle A_i,

$$\boxed{\bar{\epsilon}_i = \int_{a_{2i-1}}^{a_{2i}} dx\, \bar{\rho}(x) = \frac{1}{2\pi i} \oint_{A_i} dx\, \bar{\omega}(x)}\,, \tag{3.53}$$

where we used eq. (3.30). The filling fraction $\bar{\epsilon}_i$ agrees with the large N limit of the fraction ϵ_i of eigenvalues that belong to the cut (a_{2i-1}, a_{2i}), and therefore with our earlier definition of filling fractions in the context of normal matrix ensembles.

Similarly, let B_i be a closed cycle on the Riemann surface that goes from a_{2i} to a_{2i+1} in the first sheet, and comes back to a_{2i} in the second sheet:



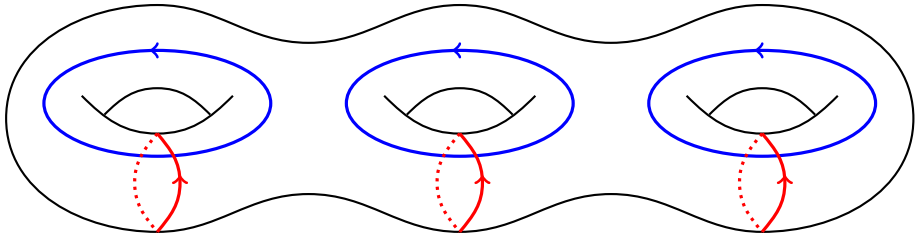
Then, the equation $\int_{a_{2i}}^{a_{2i+1}} V'_{\text{eff}} = 0$ (3.45) can be rewritten as

$$\boxed{\oint_{B_i} dx\, \bar{\omega}(x) = 0}\ , \tag{3.54}$$

where we used eq. (3.44), which implies that V'_{eff} changes sign when we change sheets, together with $\oint_{B_i} V' = 0$. The integrals (3.53) and (3.54) of the one-form $\bar{\omega}(x)$ on the cycles of our Riemann surface are called the **periods**.

Cycles

A compact Riemann surface Σ of genus g has $2g$ independent non-contractible cycles:



The set of integer linear combinations of non-contractible cycles is the integer homology space

$$H_1(\Sigma, \mathbb{Z}) = \bigoplus_{i=1}^{g} (\mathbb{Z}A_i \oplus \mathbb{Z}B_i)\ , \tag{3.55}$$

where $\{A_i, B_i\}_{i=1,\dots,g}$ is a set of independent cycles. There always exist a symplectic basis of cycles, that is a basis such that

$$A_i \cap B_j = \delta_{i,j}\,, \quad A_i \cap A_j = \emptyset\,, \quad B_i \cap B_j = \emptyset\,. \tag{3.56}$$

Period matrix

We have introduced periods as integrals of the one-form $\bar{\omega}(x)dx$. Let us now use other one-forms and consider the more general periods

$$\eta_{k,i} = \frac{1}{2\pi i} \oint_{A_i} x^{k-1} \frac{dx}{\sqrt{\sigma(x)}} \quad , \quad \tilde{\eta}_{k,i} = \frac{1}{2\pi i} \oint_{B_i} x^{k-1} \frac{dx}{\sqrt{\sigma(x)}}\ . \tag{3.57}$$

It can be proved that the square matrix $[\eta_{k,i}]_{i,k=1,\dots,g}$ is invertible, and we define the **period matrix** of our Riemann surface as

$$\tau = \eta^{-1}\tilde{\eta}\ . \tag{3.58}$$

It can be proved that τ belongs to the **Siegel space** of genus g: the set of size g symmetric complex matrices, whose imaginary part is positive definite. However, in the Siegel space of real dimension $\frac{1}{2}g(g+1)$, only a submanifold of dimension g corresponds to period matrices of Riemann surfaces. That submanifold parametrizes the complex structure moduli of genus g Riemann surfaces. These moduli correspond to the coefficients of the polynomial $\bar{P}(x)$ that are not fixed by the constraint that $(V')^2 - 4\bar{P}$ has $d - s$ double zeros.

3.3.2 Conformal maps and universality

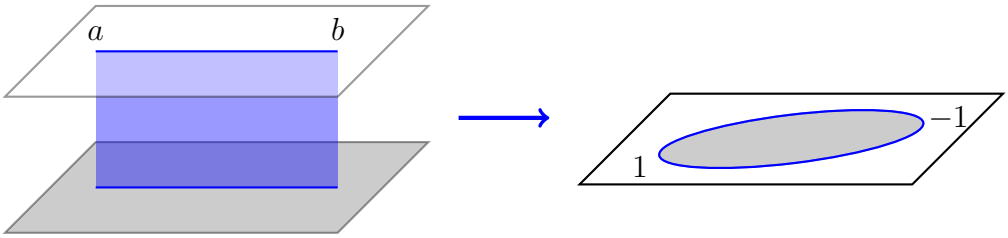
We now demonstrate how the spectral curve allows us to efficiently compute large N correlation functions. The idea is that all Riemann surfaces of the same genus and with the same complex structure are conformally equivalent. For example, any Riemann surface of genus zero has an invertible conformal map to the Riemann sphere. Using this conformal map simplifies calculations, and leads to universal results.

Genus zero: the Joukowsky map

We start with the genus zero case, where matrix eigenvalues live in a two-sheeted complex plane with one cut $x \in [a,b]$. This is mapped to the Riemann sphere $\bar{\mathbb{C}}$ with coordinate z by the **Joukowsky map**,

$$\boxed{x = \frac{a+b}{2} + \gamma\left(z + \frac{1}{z}\right) \quad \text{with} \quad \gamma = \frac{a-b}{4}} \,. \tag{3.59}$$

In particular, the cut is mapped to the unit circle $|z| = 1$,



This map was originally introduced for simplifying the study of the flow of air around an airplane's wing, by mapping the wing's section to a circle.

Let us rewrite the resolvent $\bar{\omega}$ (3.36) in terms of the Joukowsky coordinate. The non-analytic contribution $\sqrt{\sigma(x)}$ becomes a rational function of z,

$$\sqrt{\sigma(x)} = \sqrt{(x-a)(x-b)} = \gamma\left(z - \frac{1}{z}\right) \,, \tag{3.60}$$

and $\bar{\omega}(z)$ must be a rational function with poles at $z = 0$ and $z = \infty$, that is $\bar{\omega}(z) \in \mathbb{C}[z, \frac{1}{z}]$. Actually, the behaviour (3.29) of $\bar{\omega}(x)$ near $x = \infty$ implies $\bar{\omega}(z) \in \mathbb{C}[\frac{1}{z}]$, and

$$\bar{\omega}(z) = \sum_{k=0}^{d} v_k z^{-k} \quad \text{with} \quad \begin{cases} v_0 = 0 \,, \\ v_1 = \frac{1}{\gamma} \,. \end{cases} \tag{3.61}$$

Changing sheets while keeping x fixed now amounts to $z \mapsto \frac{1}{z}$, and the Riemann–Hilbert equation (3.27) becomes

$$V'(x) = \bar{\omega}(z) + \bar{\omega}\left(\frac{1}{z}\right) \,. \tag{3.62}$$

This determines v_k in terms of a, b and V'. The equations (3.61) for v_0 and v_1 then determine a and b.

For example, consider the case of an even, quartic potential

$$V(x) = \frac{1}{t}\left(\frac{x^2}{2} - \frac{x^4}{4}\right) \quad \text{with} \quad t < 0 \,. \tag{3.63}$$

By symmetry we assume that $b = -a$, and thus $(a, b) = (2\gamma, -2\gamma)$. We then find $(v_0, v_1, v_2, v_3) = \left(0, \frac{1}{t}(\gamma - 3\gamma^3), 0, -\frac{\gamma^3}{t}\right)$. The condition $v_1 = \frac{1}{\gamma}$ implies $\gamma^2 = \frac{1}{6}(1 + \sqrt{1-12t})$, which determines the position of the cut in terms of the parameter t. Using $\frac{1}{z} = \frac{x - \sqrt{x^2-4\gamma^2}}{2\gamma}$, we then compute

$$\bar{\omega}(x) = \frac{1}{\gamma z} - \frac{\gamma^3}{t z^3} = \frac{1}{2t}\left(x - x^3 + (x^2 - 1 + 2\gamma^2)\sqrt{x^2 - 4\gamma^2}\right) , \tag{3.64}$$

$$\bar{\rho}(x) = \frac{1}{2\pi t}(x^2 - 1 + 2\gamma^2)\sqrt{4\gamma^2 - x^2} , \tag{3.65}$$

where $\bar{\rho}$ is deduced from $\bar{\omega}$ via eq. (3.30).

Our one-cut treatment of the quartic potential is however valid only provided the resulting equilibrium density $\bar{\rho}$ is positive on its support $(2\gamma, -2\gamma)$, and therefore provided $\gamma^2 \geq \frac{1}{2}$ i.e. $t \leq -\frac{1}{4}$. For $t > -\frac{1}{4}$ we must introduce two cuts, and therefore a spectral curve with the genus $g = 1$. By symmetry, we look for cuts of the form $(-a, -b) \cup (b, a)$ and for an odd polynomial $M(x) = -\lambda x$. Using eq. (3.36), we have

$$\bar{\omega}(x) = \frac{1}{2t}\left(x - x^3 + \lambda x\sqrt{(x^2 - a^2)(x^2 - b^2)}\right) . \tag{3.66}$$

The values of λ and of the spectral edges a, b can be deduced from the asymptotic behaviour of $\bar{\omega}(x)$ (3.29),

$$\lambda = 1 \quad , \quad \begin{cases} a^2 = 1 + 2\sqrt{-t} , \\ b^2 = 1 - 2\sqrt{-t} . \end{cases} \tag{3.67}$$

For any $t \in (-\frac{1}{4}, 0)$ we have $a^2 > b^2 > 0$, and $\bar{\rho}$ is positive on its support.

Genus one: elliptic functions

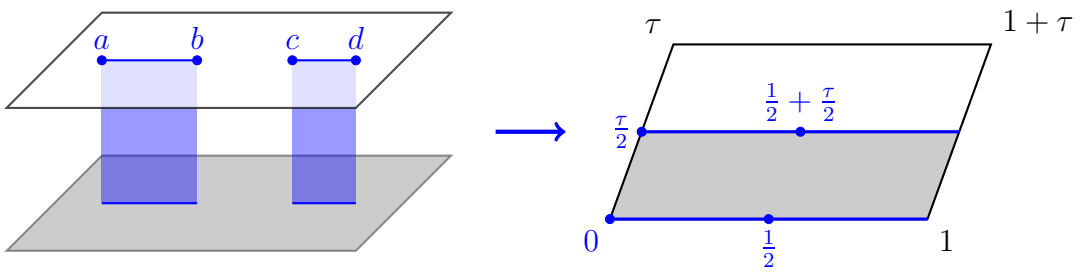
The double-covered complex plane with two cuts (a, b) and (c, d) is conformally equivalent to a torus, which we realize as the parallelogram $\frac{\mathbb{C}}{\mathbb{Z} + \tau\mathbb{Z}}$. The coordinate u on the torus is given in terms of the coordinate x on the complex plane by

$$u(x) = \frac{1}{2\int_a^b \frac{dx'}{\sqrt{\sigma(x')}}} \int_a^x \frac{dx'}{\sqrt{\sigma(x')}} \quad \text{with} \quad \sigma(x) = (x-a)(x-b)(x-c)(x-d) , \tag{3.68}$$

where different choices of integration contours lead to different representations of $u \equiv u + 1 \equiv u + \tau$. We have $u(a, b, c, d) = (0, \frac{1}{2}, \frac{\tau}{2}, \frac{1}{2} + \frac{\tau}{2})$, and the complex structure modulus of the torus is

$$\tau = 2u(c) = i\frac{F(\frac{1}{2}, \frac{1}{2}, 1, 1 - r)}{F(\frac{1}{2}, \frac{1}{2}, 1, r)} \quad \text{with} \quad r = \frac{(a-b)(c-d)}{(a-d)(c-b)} , \tag{3.69}$$

where r is called the cross-ratio of the four points (a, b, c, d), and F is the hypergeometric function.



The resolvent $\bar{\omega}$ is then not only a meromorphic, but also an **elliptic function** of u, that is a doubly periodic function such that

$$\bar{\omega}(u+1) = \bar{\omega}(u+\tau) = \bar{\omega}(u) \ . \tag{3.70}$$

Such a meromorphic, elliptic function is entirely characterized by its behaviour at its poles.

The map $u(x)$ (3.68) can be rewritten in terms of standard special functions, provided the spectral edges (a, b, c, d) have special positions in the complex plane. Such special positions can always be reached from generic positions via a Möbius transformation, at the expense of changing the potential $V'(x)$ from a polynomial to a rational function of x. A first type of special positions is $(a, b, c, d) = (-\frac{1}{k}, -1, 1, \frac{1}{k})$, in which case x and $\sigma(x)$ can be written in terms of the Jacobi elliptic functions,

$$x = \mathrm{sn}(u,k) \quad , \quad \sqrt{\sigma(x)} = \mathrm{cn}(u,k)\,\mathrm{dn}(u,k) \ . \tag{3.71}$$

Another type of special positions is $d = \infty$ with $a + b + c = 0$, in which case x and $\sigma(x)$ can be written in terms of a Weierstrass elliptic $\wp$-function,

$$x = \wp(u) \quad , \quad \sqrt{\sigma(x)} = \frac{1}{2}\wp'(u) \ . \tag{3.72}$$

This elliptic function obeys the differential equation

$$(\wp'(u))^2 = 4\sigma(\wp(u)) = 4\wp(u)^3 - g_2\wp(u) - g_3 \ , \tag{3.73}$$

where the parameters g_2, g_3 are given in terms of the roots a, b, c of $\sigma(x)$ by

$$g_2 = 4(ab + ac + bc) \quad , \quad g_3 = -4abc \ . \tag{3.74}$$

Higher genus

A Riemann surface Σ of genus $g \geq 2$ can be mapped to a polygon with $4g$ sides in the hyperbolic plane $\mathcal{H}$, whose sides are glued pairwise. That polygon is a model of the quotient of $\mathcal{H}$ by a discrete group Γ of hyperbolic isometries, and

$$\Sigma \sim \frac{\mathcal{H}}{\Gamma} \ . \tag{3.75}$$

All meromorphic functions on $\frac{\mathcal{H}}{\Gamma}$ can be written in terms of theta functions – higher genus generalizations of elliptic functions [47].

3.3.3 The two-point function

In the large N limit, the non-connected two-point function $\rho_2(x_1, x_2) + N^2\rho(x_1)\rho(x_2)$ is dominated by the trivial term $N^2\bar{\rho}(x_1)\bar{\rho}(x_2)$. The large N limit of the connected two-point function $\bar{\rho}_2(x_1, x_2) = \lim_{N\to\infty} \rho_2(x_1, x_2)$ appears as a subleading contribution. While we do not a priori expect the saddle point approximation to be useful for computing such subleading terms, we can still obtain some useful information.

Let us renounce the large N limit for a moment, and come back to our original eigenvalue integral (3.3). Using the functional derivative $\frac{\delta}{\delta V(x)}$ with respect to the potential $V(x)$, which in particular obeys $\frac{\delta V(x')}{\delta V(x)} = \delta(x - x')$, we can write the density of eigenvalues as

$$\rho(x) = -\frac{1}{N^2}\frac{\delta \log \mathcal{Z}}{\delta V(x)} \ . \tag{3.76}$$

Taking an extra functional derivative, we obtain

$$\frac{\delta\rho(x)}{\delta V(x')} = -\rho_2(x,x') \ . \tag{3.77}$$

We now want to take the large N limit of this exact relation. The problem is that the large N limit in general does not commute with the functional derivative: in particular, the limit of $\rho_2(x,x')$ can receive contributions from the subleading terms of $\rho(x)$, which have oscillations of frequency $O(N)$. (See fig. (3.20).) However, in the one-cut case, these oscillations are absent, and we can take the large N limit of our functional derivation relation. Taking in addition the double Stieltjes transform, we obtain

$$\int_{\operatorname{supp}\bar\rho} \frac{dx''}{x'-x''} \frac{\delta\bar\omega(x)}{\delta V(x'')} = -\bar\omega_2(x,x') \ , \tag{3.78}$$

where $\bar\omega_2$ is the double Stieltjes transform of $\bar\rho_2$,

$$\bar\omega_2(x,x') = \int_{\operatorname{supp}\bar\rho}\int_{\operatorname{supp}\bar\rho} \frac{d\lambda}{x-\lambda}\frac{d\lambda'}{x'-\lambda'}\bar\rho_2(\lambda,\lambda') \ . \tag{3.79}$$

We can now obtain a Riemann–Hilbert equation for $\bar\omega_2$ by applying the functional derivative to the Riemann–Hilbert equation (3.27) for $\bar\omega$,

$$\boxed{\bar\omega_2(x+i0,x') + \bar\omega_2(x-i0,x') = -\frac{1}{(x-x')^2}} \ . \tag{3.80}$$

This means that $\bar\omega_2(x,x') + \frac{1}{2(x-x')^2}$ changes sign when x goes through a cut, and therefore behaves just like $\sqrt{\sigma(x)}$ in this respect. This implies

$$\bar\omega_2(x,x') = -\frac{1}{2(x-x')^2}\left(1 - \frac{Q_2(x,x')}{\sqrt{\sigma(x)}\sqrt{\sigma(x')}}\right) \ , \tag{3.81}$$

where $Q_2(x,x')$ is meromorphic in its two arguments. Let us analyze how $Q_2(x,x')$ behaves near its possible singularities:

- Near $x = x'$, the regularity of $\bar\omega_2(x,x')$ implies $\lim\limits_{x\to x'} Q_2(x,x') = \sigma(x)$ as well as $\lim\limits_{x\to x'} \partial_x Q_2(x,x') = \frac{1}{2}\sigma'(x)$.
- Near a spectral edge $x = a$, the worst possible singularity of $\bar\omega(x)$ is $O(\sqrt{x-a})$. Applying $\frac{\delta}{\delta V(x')} = \frac{\delta}{\delta a}\frac{\delta a}{\delta V(x')}$, we obtain an $O(\frac{1}{\sqrt{x-a}})$ singularity for $\bar\omega_2(x,x')$. This shows that $Q(x,x')$ must be regular at a, and therefore Q_2 must be a polynomial.
- Near $x = \infty$, we must have by definition $\bar\omega_2(x,x') \underset{x\to\infty}{=} O(\frac{1}{x^2})$, which implies $\deg_x Q \leq \frac{1}{2}\deg\sigma = s$. Using again the behaviour near $x = x'$, we actually find $\deg_x Q = s$.

In the case where we have only one cut (a,b), these constraints fully determine $Q_2(x,x')$, and we find

$$Q_2(x,x') = xx' - \frac{a+b}{2}(x+x') + ab \ . \tag{3.82}$$

Using the Joukowsky variables z_1, z_2 which correspond to x_1, x_2 (3.59), this leads to

$$\bar{\omega}_2(x_1, x_2)dx_1dx_2 = \frac{dz_1dz_2}{(z_1 - z_2)^2} - \frac{dx_1dx_2}{(x_1 - x_2)^2} \ . \tag{3.83}$$

We can then compute the connected two-point function $\bar{\rho}_2$ as the discontinuity of $\bar{\omega}_2$ on the cut. Notice that the second term of $\bar{\omega}_2(x_1, x_2)$ is a rational fraction of x_1, x_2, therefore it is continuous on the cut and does not contribute to the two-point function. While that term ensures that $\bar{\omega}_2(x_1, x_2)$ is regular at $x_1 = x_2$, all the physical information is contained in the term

$$\boxed{B(z_1, z_2) = \frac{dz_1dz_2}{(z_1 - z_2)^2} = d_{z_1}d_{z_2}\log(z_1 - z_2)} \ . \tag{3.84}$$

This two-form is called the **fundamental second kind differential** of the Riemann sphere. It is manifestly related to the Green's function $G(z_1, z_2) = \log|z_1 - z_2|$ of the Laplacian $\Delta = 4\partial\bar{\partial}$, which is the real, symmetric solution of

$$\Delta_{z_1}G(z_1, z_2) = 2\pi\delta(z_1 - z_2) \ . \tag{3.85}$$

We can actually rewrite the two-point function in terms of the fundamental second kind differential as

$$\boxed{\bar{\omega}_2(x_1, x_2)dx_1dx_2 = -B\left(z_1, \frac{1}{z_2}\right)} \ , \tag{3.86}$$

using the identity

$$\frac{dxdx'}{(x - x')^2} = B(z, z') + B\left(z, \frac{1}{z'}\right) \ , \tag{3.87}$$

which itself follows from the Joukowsky map.

The expression for the two-point function in terms of the fundamental second kind differential, which we just saw in the one-cut case of the Riemann sphere, has a generalization to the multi-cut case of an arbitrary Riemann surface. The fundamental second kind differential B of a Riemann surface is defined as a meromorphic symmetric bilinear differential whose only singularity is a double pole on the diagonal. It is unique up to holomorphic terms and a choice of normalization, and only depends on the complex structure of the Riemann surface. As we will see in Section 4.5, the large N two-point function is not given solely by the fundamental second kind differential, but may also involve oscillatory terms. The resulting expression is more complicated than in the one-cut case, but it is still universal.

3.4 Examples

3.4.1 Generalized probability laws

The $O(n)$ matrix model

The partition function of the $O(n)$ matrix model is proportional to the eigenvalue integral

$$\mathcal{Z} \propto \int_{\mathbb{R}^N} \prod_i d\lambda_i \, e^{-NV(\lambda_i)} \frac{\prod_{i<j}(\lambda_i - \lambda_j)^2}{\prod_{i,j}(\lambda_i + \lambda_j)^{\frac{n}{2}}} \ , \tag{3.88}$$

where we have done the rescaling $V \to NV$ of the potential. The saddle point equation is

$$\forall x \in \operatorname{supp} \bar{\rho}\,, \qquad V(x) - 2 \fint dx'\, \bar{\rho}(x') \log|x-x'| + n \int dx'\, \bar{\rho}(x') \log|x+x'| = \ell\,. \quad (3.89)$$

(Compare with the saddle point equation (3.25) of our Hermitian matrix model.) The corresponding Riemann–Hilbert equation is

$$\forall x \in \operatorname{supp} \bar{\rho}\,, \qquad V'(x) = \bar{\omega}(x+i0) + \bar{\omega}(x-i0) + n\bar{\omega}(-x)\,. \quad (3.90)$$

This equation can in general be solved using theta functions. If however $n \in 2\cos \pi\mathbb{Q}$, then the solution $\bar{\omega}(x)$ is an algebraic function.

More general interactions between eigenvalues

The partition function of the $O(n)$ matrix model is an example of a matrix integral of the type

$$\mathcal{Z} \propto \int_{\mathbb{R}^N} \prod_i e^{-NV(\lambda_i)} d\lambda_i \prod_{i<j} (\lambda_i - \lambda_j)^2 R(\lambda_i, \lambda_j)\ , \quad (3.91)$$

where $R(x,x') = R(x',x)$ is an analytic function such that $R(x,x) \neq 0$. The corresponding Riemann–Hilbert equation is

$$\forall x \in \operatorname{supp} \bar{\rho}\,, \quad V'(x) = \bar{\omega}(x+i0) + \bar{\omega}(x-i0) + \frac{1}{2\pi i} \oint_{\mathcal{C}} \partial_x \log R(x,x') \bar{\omega}(x') dx'\,, \quad (3.92)$$

where the contour $\mathcal{C}$ surrounds $\operatorname{supp} \bar{\rho}$, and can be deformed to surround the zeros and poles of R.

While this leads to more complicated spectral curves than our Hermitian matrix model, the universal results continue to hold. In particular, the solutions of the Riemann–Hilbert equation are still parametrized by periods of the spectral curve, and correspond to possible integration contours. The two-point function is still the fundamental second kind differential of the spectral curve, plus possible oscillatory terms.

An application to knot theory

An interesting example is the matrix model with the probability law (1.69), which computes the HOMFLY polynomials of torus knots. In the case of the torus knot $\mathfrak{K}_{P,Q}$, the relevant eigenvalue integral is

$$\mathcal{Z} \propto \int_{\mathbb{R}^N} \prod_{i<j} 2\sinh\left(\frac{\lambda_i - \lambda_j}{2P}\right) 2\sinh\left(\frac{\lambda_i - \lambda_j}{2Q}\right) \prod_{i=1}^N e^{-\frac{N}{2PQt}\lambda_i^2} d\lambda_i\ , \quad (3.93)$$

$$= \int_{\mathbb{R}_+^N} \prod_{i<j} (x_i^Q - x_j^Q)(x_i^P - x_j^P) \prod_{i=1}^N e^{-\frac{N}{2PQt}(\log x_i)^2} e^{\log x_i} dx_i\ , \quad (3.94)$$

where $\lambda_i = PQ \log x_i$. This corresponds to the potential and interaction function

$$V(x) = \frac{1}{2PQt} \log x^2 - \frac{1}{N} \log x \quad , \quad R(x,x') = \frac{(x^P - x'^P)(x^Q - x'^Q)}{(x-x')^2}\ . \quad (3.95)$$

In the saddle point approximation, we neglect the $O(\frac{1}{N})$ term in the potential, and we obtain the Riemann–Hilbert equation

$$\frac{1}{PQt}\frac{\log x}{x} = \bar{\omega}(x+i0)+\bar{\omega}(x-i0)+\sum_{k=1}^{P-1} e^{\frac{2\pi ik}{P}}\bar{\omega}(e^{\frac{2\pi ik}{P}}x)+\sum_{k=1}^{Q-1} e^{\frac{2\pi ik}{Q}}\bar{\omega}(e^{\frac{2\pi ik}{Q}}x)\ . \quad (3.96)$$

This is not an algebraic equation for the function $\bar{\omega}$, as several different values of its argument appear. However, this leads to an algebraic equation for the function

$$f(x)=\prod_{k=1}^{P-1} y(e^{\frac{2\pi ik}{P}}x) \qquad \text{with} \qquad y(x)=-e^{-\frac{P+Q}{2PQ}t}xe^{-\frac{P+Q}{PQ}x\bar{\omega}(x)}\ . \quad (3.97)$$

The function $y(x)$ indeed satisfies

$$y(x+i0)y(x-i0)\prod_{k=1}^{P-1} y(e^{\frac{2\pi ik}{P}}x)\prod_{k=1}^{Q-1} y(e^{\frac{2\pi ik}{Q}}x)=1\ , \quad (3.98)$$

which then implies that f satisfies a polynomial equation of degree $P+Q$ of the form

$$0=1+f^{P+Q}+(-1)^{Q+PQ}f^Q x^{-PQ}e^{\frac{P+Q}{2}t}+(-1)^{P+PQ}f^P x^{PQ}e^{\frac{P+Q}{2}t}+\sum_{k=1}^{P+Q-1} s_k f^k\ . \quad (3.99)$$

The $P+Q-1$ unknown coefficients s_k play the same role as the polynomial $\bar{P}(x)$ in the case of our Hermitian matrix model. In particular, the assumption that we have only one cut fixes all these coefficient, and f has the following rational parametrization in terms of $z\in\bar{\mathbb{C}}$,

$$\begin{cases} x^{PQ}=e^{t\frac{P+Q}{2}}\,z^{-P}\left(\frac{1-e^{-t}z}{1-z}\right)^Q\ , \\ f=-z\frac{1-e^{-t}z}{1-z}\ . \end{cases} \quad (3.100)$$

Using the equilibrium density that is determined by this saddle point computation, it is possible to find the leading asymptotic behaviour of the HOMFLY polynomial of the torus knot $\mathfrak{K}_{P,Q}$.

3.4.2 Multi-matrix models

Let us consider the matrix chain with the partition function

$$\mathcal{Z}=\prod_{j=1}^{k}\int_\Gamma dM_j\ e^{-N\,\mathrm{Tr}\,V_j(M_j)}e^{N\,\mathrm{Tr}\,M_jM_{j+1}}\ , \quad (3.101)$$

where $V_1,\dots V_k$ are polynomials. In order for a saddle point approximation to exist, the potentials for the integrated matrix variables M_j come with a prefactor N, just as in the one-matrix model. Moreover, we assume that the external field M_{k+1} has an N-independent number of eigenvalues m, with the multiplicity of the eigenvalue λ_i behaving as

$$\#\lambda_i \underset{N\to\infty}{=} N\nu_i+o(N)\ , \quad (3.102)$$

where ν_i is N-independent. Assuming that the potentials V_j are polynomial, the Stieltjes transform $\bar{\omega}$ of the equilibrium density of eigenvalues of the first matrix M_1 is then an algebraic function [48], that we will now describe. We will start with two special cases before dealing with the general case.

One-matrix model with an external field

In the case $k = 1$, the function $\bar{\omega}$ should obey a generalization of the large N saddle point equation (3.18) for the one-matrix model without an external field. A prominent role in that equation is played by the polynomial $\bar{P}(x)$, and it turns out that the correct generalization of that polynomial is a function of two variables,

$$\bar{P}(x,y) = \lim_{N\to\infty} \frac{1}{N} \left\langle \mathrm{Tr}\, \frac{V_1'(x) - V_1'(M_1)}{x - M_1} \frac{1}{y - M_2} \right\rangle \ . \tag{3.103}$$

This is a polynomial of x of degree $\deg V_1' - 1$, and a rational function of y of degree m with simple poles at $y = \lambda_i$, such that

$$\bar{P}(x,y) \underset{x\to\infty}{\sim} \frac{V'(x)}{x} \sum_{i=1}^{m} \frac{\nu_i}{y - \lambda_i} \ . \tag{3.104}$$

Defining moreover

$$Y_2(x) = V_1'(x) - \bar{\omega}(x) \ , \tag{3.105}$$

the large N saddle point equation is

$$V_1'(x) - Y_2(x) = \bar{P}(x, Y_2(x)) \ , \tag{3.106}$$

which reduces to eq. (3.18) in the case $M_2 = 0$. So Y_2 and therefore also $\bar{\omega}$ obey a rational equation. Multiplying both sides with the denominator, this is equivalent to a polynomial equation, and $\bar{\omega}$ is an algebraic function of x.

For example, let us consider the case of the **Kontsevich integral**, which is obtained for a cubic potential $V_1(x) = \frac{1}{3}x^3$. The polynomial $\bar{P}(x,y)$ is of the type

$$\bar{P}(x,y) = 2xR(y) + Q(y) \quad \text{with} \quad \begin{cases} R(y) = \frac{1}{2}\sum_{i=1}^{m} \frac{\nu_i}{y-\lambda_i} \ , \\ Q(y) = \sum_{i=1}^{N} \frac{\nu_i c_i}{y-\lambda_i} \ , \end{cases} \tag{3.107}$$

where the m unknown coefficients c_i are determined either by eq. (3.54), or (if we have fixed filling fractions) by eq. (3.53). The large N saddle point equation is

$$x^2 - 2R(Y_2)x - Q(Y_2) - Y_2 = 0 \ . \tag{3.108}$$

This amounts to an algebraic equation of degree $m+1$ in Y_2 and of degree 2 in x. So $x(Y_2)$ has two branches, and can easily be computed,

$$x_\pm(Y_2) = R(Y_2) \pm \sqrt{R(Y_2)^2 + Y_2 + Q(Y_2)} \ . \tag{3.109}$$

In the case where the corresponding algebraic curve has genus zero, let us look for a rational parametrization of that curve, in other words for an analogue of the Joukowsky map. We want to express both x and Y_2 as functions of a variable z on the Riemann sphere $\bar{\mathbb{C}}$. Assuming $Y_2(\infty) = \infty$, and calling $z = \zeta_i$ the positions of the poles, we have This leads to

$$x(z) = z + \sum_{i=1}^{m} \frac{\nu_i}{2\zeta_i(z - \zeta_i)} \ , \tag{3.110}$$

$$Y_2(z) = z^2 + \sum_{i=1}^{m} \frac{\nu_i}{\zeta_i} \ , \tag{3.111}$$

which turns out to be the unique rational parametrization of our algebraic curve, modulo Möbius transformations of z. The parameters ζ_i are determined by $Y_2(\zeta_i) = \lambda_i$, and we moreover have the condition

$$c_i = -x(-\zeta_i) = \zeta_i + \sum_{j=1}^{m} \frac{\nu_j}{2\zeta_j(\zeta_i + \zeta_j)} \ . \tag{3.112}$$

So the assumption that our spectral curve has genus zero completely determines the parameters c_i.

Two-matrix model

Let us consider the case $k = 2$ in the absence of an external field, equivalently with $M_3 = 0$. We now define

$$\bar{P}_2(x,y) = \lim_{N\to\infty} \frac{1}{N} \left\langle \mathrm{Tr} \frac{V_1'(x) - V_1'(M_1)}{x - M_1} \frac{V_2'(y) - V_2'(M_2)}{y - M_2} \right\rangle - 1 \ , \tag{3.113}$$

which is a polynomial of x of degree $\deg V_1' - 1$, and a polynomial of y of degree $\deg V_2' - 1$. The large N saddle point equation for the resolvent $\bar{\omega}$ turns out to be

$$(V_1'(x) - Y_2(x))(V_2'(Y_2(x)) - x) = \bar{P}_2(x, Y_2(x)) \ , \tag{3.114}$$

where $Y_2(x)$ is still given by eq. (3.105). This is a polynomial equation of degree $\deg V_1$ in x and $\deg V_2$ in Y_2, whose solution $\bar{\omega}$ is therefore an algebraic function of x. The unknown coefficients of $\bar{P}_2(x,y)$ are determined by equations of the type (3.54), or (3.53) if we have fixed filling fractions.

In the case where the corresponding algebraic curve has genus zero, that curve has a rational parametrization of the type

$$x(z) = \gamma z + \sum_{k=0}^{\deg V_2'} \alpha_k z^{-k} \ , \tag{3.115}$$

$$Y_2(z) = \frac{\gamma}{z} + \sum_{k=0}^{\deg V_1'} \beta_k z^k \ , \tag{3.116}$$

where the coefficients γ, α_k, β_k, and the unknown coefficients of $\bar{P}_2(x,y)$, can be deduced from our polynomial equation. To do this, let us study the asymptotics of that equation at $z = 0, \infty$. Using $\bar{P}_2(x,y) \underset{x,y\to\infty}{\sim} \frac{V_1'(x)}{x} \frac{V_2'(y)}{y}$, we find

$$V_1'(x(z)) - Y_2(z) \underset{z\to\infty}{=} \frac{1}{\gamma z} + O\left(\frac{1}{z^2}\right) \quad , \quad V_2'(Y_2(z)) - x(z) \underset{z\to 0}{=} \frac{z}{\gamma} + O(z^2) \ . \tag{3.117}$$

Actually, the coefficients of the leading terms of these two expressions are left undetermined by the polynomial equation. To find the leading coefficient of the first expression, we need the asymptotic behaviour (3.29) of $\bar{\omega}$, together with eq. (3.105). We then already have enough equations for determining the $\deg V_1' + \deg V_2' + 3$ unknown coefficients $\gamma, \alpha_k, \beta_k$. Nevertheless, we can easily determine the leading coefficient of the second expression, by relating it to the leading coefficient of the first expression using $\oint_{z=\infty} Y_2 dx = -\oint_{z=0} Y_2 dx = \oint_{z=0} x dY_2$.

Matrix chain with an external field

We now consider the most general multi-matrix model that obeys our assumptions. Let us generalize our definition (3.105) of $Y_2(x)$, and define functions $Y_0(x), \dots, Y_{k+1}(x)$ by

$$Y_0(x) = \bar{\omega}(x) \quad , \quad Y_1(x) = x \quad , \quad Y_{j+1}(x) = V_j'(Y_j(x)) - Y_{j-1}(x) \ . \tag{3.118}$$

All these functions are polynomials of x and $\bar{\omega}(x)$. We now define

$$\bar{P}(y_1, \dots, y_{k+1}) = \lim_{N\to\infty} \frac{1}{N} \left\langle \operatorname{Tr} \operatorname{Pol}_{y_1,\dots,y_k} f(y_1, \dots, y_k) \frac{1}{y_1 - M_1} \cdots \frac{1}{y_{k+1} - M_{k+1}} \right\rangle \ , \tag{3.119}$$

where

$$f(y_1, y_2, \dots, y_k) = \det \begin{pmatrix} V_1'(y_1) & -y_2 & & & \\ -y_1 & V_2'(y_2) & -y_3 & & \\ & & \ddots & & \\ & & -y_{k-2} & V_{k-1}'(y_{k-1}) & -y_k \\ & & & -y_{k-1} & V_k'(y_k) \end{pmatrix} \ , \tag{3.120}$$

and $\operatorname{Pol}_x g(x)$ is the polynomial part of the Taylor expansion of $g(x)$ near $x = \infty$. In particular $\operatorname{Pol}_x \frac{V'(x)}{x - M_1} = \frac{V'(x) - V'(M_1)}{x - M_1}$, and we recover the polynomial $\bar{P}(x, y)$ (3.103) of the one-matrix model with an external field as a special case. In the case of the two-matrix model, we recover the appropriate polynomial $\bar{P}_2(x, y)$ up to an overall factor, $\bar{P}(y_1, y_2, y_3) = \frac{1}{y_3} \bar{P}_2(y_1, y_2)$.

The large N saddle point equations then reduce to a system of algebraic equations,

$$Y_0 = \bar{P}(Y_1, \dots, Y_{k+1}) \quad , \quad Y_{j-1} + Y_{j+1} = V_j'(Y_j) \quad , \quad (1 \leq j \leq k) \ . \tag{3.121}$$

In the case of the two-matrix model, we obtain the equations

$$Y_0 Y_3 = \bar{P}(Y_1, Y_2) \quad , \quad Y_0 + Y_2 = V_1'(Y_1) \quad , \quad Y_1 + Y_3 = V_2'(Y_2) \ , \tag{3.122}$$

which are equivalent to (3.114).

The function $\bar{P}(y_1, \dots, y_{k+1})$ is a polynomial of $y_1, \dots, y_k$ of degree at most $\deg V_j' - 1$ in y_j, and a rational fraction of the y_{k+1}, with m simple poles at the eigenvalues of the external field M_{k+1}. The total number of unknown coefficients of $\bar{P}$ is therefore

$$d = m \prod_{j=1}^{k} \deg V_j' \ . \tag{3.123}$$

Actually, one of these coefficients is fixed by the asymptotic behaviour (3.29) of the resolvent $\bar{\omega}(x)$. The remaining $d - 1$ coefficients are in one-to-one correspondence with the periods of the spectral curve, and also with the choices of integration contours for the random matrices. In contrast to the case of the one-matrix model with no external field, the spectral curve can have more than two sheets, in other words the degree of its equation in x can be larger than two.

In the case where the spectral curve has genus zero, there must exist a rational parametrization of that curve, in other words an analogue of the Joukowsky map. This parametrization is given by functions $Y_j(z)$ with $z \in \bar{\mathbb{C}}$ that obey the algebraic equations

(3.121). Moreover, the asymptotic behaviour of $\bar{\omega}(x) = Y_0(Y_1)$ near $x = \infty$ implies that, under the assumption $\lim_{z\to\infty} Y_1(z) = \infty$, we must have

$$Y_0(z) \underset{z\to\infty}{=} \frac{1}{Y_1(z)} + O\left(\frac{1}{Y_1(z)^2}\right) . \tag{3.124}$$

And our assumptions on the external field M_{k+1} imply that near a point ζ_i such that $Y_{k+1}(\zeta_i) = \lambda_i$, we must have

$$Y_k(z) \underset{z\to\zeta_i}{=} \frac{\nu_i}{Y_{k+1}(z) - \lambda_i} + O(1) . \tag{3.125}$$

To prove this, assume $\deg V_j' \geq 2$, and notice that $\lim_{z\to\zeta_i} Y_j(z) = \infty$ for $0 \leq j \leq k$, with $Y_0 \gg Y_1 \gg \cdots \gg Y_{k+1}$. The leading behaviour of the saddle point equations (3.121) is

$$Y_0 \sim \prod_{j=1}^{k} \frac{V_j'(Y_j)}{Y_j} \cdot \frac{\nu_i}{Y_{k+1} - \lambda_i} \quad , \quad Y_{j-1} \sim V_j'(Y_j) \quad , \quad (1 \leq j \leq k) , \tag{3.126}$$

which yields the behaviour of Y_k, since all other factors in the first equation cancel with one another. More explicitly, the functions $Y_j(z)$ can be chosen as

$$Y_j(z) = \sum_{l=0}^{r_j} \alpha_{j,l} z^l + \sum_{i=1}^{m} \nu_i \sum_{l=1}^{s_j} \frac{\beta_{j,i,l}}{(z-\zeta_i)^l} , \tag{3.127}$$

where

$$r_0 = -1 , \quad r_{j\geq 1} = \prod_{l=1}^{j-1} \deg V_l' \quad , \quad s_{j\leq k} = \prod_{j+1}^{k} \deg V_l' , \quad s_{k+1} = 0 , \tag{3.128}$$

so that in particular $Y_0(z)$ has no polynomial part, while $Y_{k+1}(z)$ is purely polynomial. Up to affine transformations of z, the coefficients $\alpha_{j,l}$ and $\beta_{j,i,l}$ are uniquely determined by the last k saddle point equations (3.121), together with $Y_{k+1}(\zeta_i) = \lambda_i$ and the asymptotic behaviour of $Y_k(z)$ near the poles (3.125). (That asymptotic behaviour amounts to $\beta_{k,i,1} = \frac{1}{Y_{k+1}'(\zeta_i)}$.) The resulting rational parametrization generalizes not only the Joukowsky map, but also the parametrizations that appeared in the two special cases that we considered.

3.4.3 Complex matrix model with a harmonic potential

Let us consider a Gaussian complex matrix model, and add a harmonic contribution to the potential. The partition function is

$$\mathcal{Z} = \int_{M_N(\mathbb{C})} dM \, e^{-N\mathcal{V}(M,M^\dagger)} , \tag{3.129}$$

where total potential

$$\mathcal{V}(x,\bar{x}) = Rx\bar{x} + V(x) + \bar{V}(\bar{x}) , \tag{3.130}$$

includes a Gaussian term with a constant coefficient R, and a harmonic term $V(x)+\bar{V}(x)$, so that

$$\partial\bar{\partial}\mathcal{V}(x,\bar{x}) = R . \tag{3.131}$$

In terms of the eigenvalues λ_i of the matrix M, the partition function is

$$\mathcal{Z} \propto \int_{\mathbb{C}^N} \prod_{i<j} |\lambda_i - \lambda_j|^2 \prod_i e^{-N\left(R|\lambda_i|^2 + V(\lambda_i) + \bar{V}(\bar{\lambda}_i)\right)} d^2\lambda_i \,. \tag{3.132}$$

The corresponding saddle point equation for the equilibrium density $\bar{\rho}$ is

$$\forall x \in \operatorname{supp} \bar{\rho}\,, \qquad \mathcal{V}(x, \bar{x}) - 2 \int d^2x' \; \bar{\rho}(x') \log|x - x'| = \ell \,. \tag{3.133}$$

(Compare with the saddle point equation (3.25) of our Hermitian matrix model.) Taking the Laplacian of this saddle point equation, and using eq. (3.131), we obtain

$$\forall x \in \operatorname{supp} \bar{\rho}, \qquad \bar{\rho}(x) = \frac{R}{\pi} \,. \tag{3.134}$$

This means that the equilibrium density of eigenvalues is constant on its support.

What remains to be found is the support itself. We already know that the support must be a domain of area $\frac{\pi}{R}$. Let us parametrize this domain by the shape of its boundary curve $\partial \operatorname{supp} \bar{\rho}$, whose equation we write as

$$\bar{x} = S(x) \,. \tag{3.135}$$

The analytic function $S(x)$, which obeys

$$\bar{S} \circ S = \mathrm{Id} \;, \tag{3.136}$$

is called the **Schwarz function** of the curve $\partial \operatorname{supp} \bar{\rho}$, and can be useful for applying the Schwarz reflection principle to the domain $\operatorname{supp} \bar{\rho}$. Then the analytic Stieltjes transform of the equilibrium density,

$$\bar{\omega}(x) = \int_{\mathbb{C}} \frac{\bar{\rho}(x')}{x - x'} d^2x' \,, \tag{3.137}$$

can be rewritten in terms of the function $S(x)$, with the help of Stokes' theorem,

$$\bar{\omega}(x) = \frac{R}{\pi} \int_{\operatorname{supp} \bar{\rho}} \frac{1}{x - x'} d^2x' = \frac{R}{2\pi i} \int_{\partial \operatorname{supp} \bar{\rho}} \frac{S(x')}{x - x'} dx' \,, \tag{3.138}$$

where we used $\frac{1}{x-x'} d^2x' = \frac{1}{2i} \frac{1}{x-x'} dx' \wedge d\bar{x}' = \frac{1}{2i} d\left(\frac{\bar{x}'}{x-x'} dx'\right)$. On the other hand, taking one one derivative of the saddle point equation gives

$$\forall x \in \operatorname{supp} \bar{\rho}\,, \qquad R\bar{x} + V'(x) = \bar{\omega}(x) \;, \tag{3.139}$$

which implies in particular

$$\forall x \in \partial \operatorname{supp} \bar{\rho}\,, \qquad RS(x) + V'(x) = \bar{\omega}(x) = \frac{R}{2\pi i} \int_{\partial \operatorname{supp} \bar{\rho}} \frac{S(x')dx'}{x - x'} \,. \tag{3.140}$$

This is an affine integral equation for the Schwarz function $S(x)$. While originally derived for $x \in \partial \operatorname{supp} \bar{\rho}$, this equation must be valid outside $\operatorname{supp} \bar{\rho}$ by analyticity.

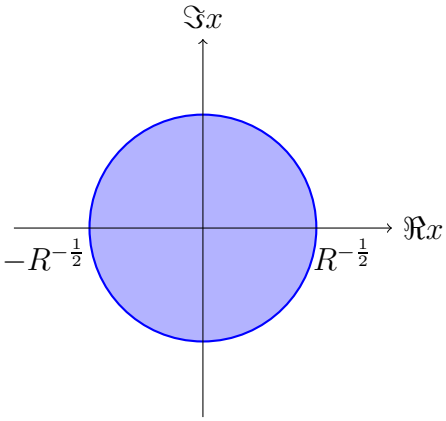
We now want to solve this equation, together with the similar equation for the function $\bar{S}$, under the constraint (3.136). Solving such an integral equation is in general difficult. As we will show in examples, in the case of polynomial or rational potentials, it is fruitful to look for solutions that obey a polynomial equation. The analysis of the behaviour of $S(x)$ as $x \to \infty$, knowing the behaviour (3.29) of $\bar{\omega}(x)$, then provides a lower bound on the degree of that polynomial, and the simplest assumption is that this bound is saturated. The coefficients of the polynomial that are not determined by the affine integral equation correspond to filling fractions.

Gaussian potential

We first consider the case $V(x) = \bar{V}(x) = 0$. From the equation (3.140) and the behaviour of $\bar{\omega}(x)$, we must have $S(x) \underset{x\to\infty}{=} \frac{1}{Rx} + O(\frac{1}{x^2})$. Repeating the analysis for $\bar{S}(x)$, and using the relation (3.136), leads to the same asymptotic behaviour for $S(x)$. This desired asymptotic behaviour is reproduced by the solutions of the degree one polynomial equation for $S(x)$,

$$S(x) - \frac{1}{Rx} = 0 \,, \tag{3.141}$$

and indeed $S(x) = \bar{S}(x) = \frac{1}{Rx}$ are solutions of our affine integral equation. The boundary of the support of the equilibrium density then has the equation $\bar{x} = \frac{1}{Rx}$, and is the circle of radius $R^{-\frac{1}{2}}$. Therefore, in the large N limit, the eigenvalues of the Gaussian matrix model fill a disk of radius $R^{-\frac{1}{2}}$,



Quadratic potentials

We consider the case $V(x) = \frac{t_2}{2}x^2$, $\bar{V}(x) = \frac{\bar{t}_2}{2}x^2$. We must have

$$S(x) \underset{x\to\infty}{=} -\frac{t_2}{R}x + \frac{1}{Rx} + O\left(\frac{1}{x^2}\right) \,, \quad \bar{S}(x) \underset{x\to\infty}{=} -\frac{\bar{t}_2}{R}x + \frac{1}{Rx} + O\left(\frac{1}{x^2}\right) \,. \tag{3.142}$$

The inverse functions therefore behave as

$$S^{-1}(x) \underset{x\to\infty}{=} -\frac{R}{t_2}x - \frac{1}{Rx} + O\left(\frac{1}{x^2}\right) \,, \quad \bar{S}^{-1}(x) \underset{x\to\infty}{=} -\frac{R}{\bar{t}_2}x - \frac{1}{Rx} + O\left(\frac{1}{x^2}\right) \,. \tag{3.143}$$

This seems at odds with the equation (3.136), which says $S = \bar{S}^{-1}$. Actually, this discrepancy means that $S(x)$ lives on a covering of the complex x-planes with at least two sheets labelled $\pm$, and the asymptotic behaviours

$$S_+(x) \underset{x\to\infty}{=} -\frac{t_2}{R}x + \frac{1}{Rx} + O\left(\frac{1}{x^2}\right) \,, \quad S_-(x) \underset{x\to\infty}{=} -\frac{R}{\bar{t}_2}x - \frac{1}{Rx} + O\left(\frac{1}{x^2}\right) \,. \tag{3.144}$$

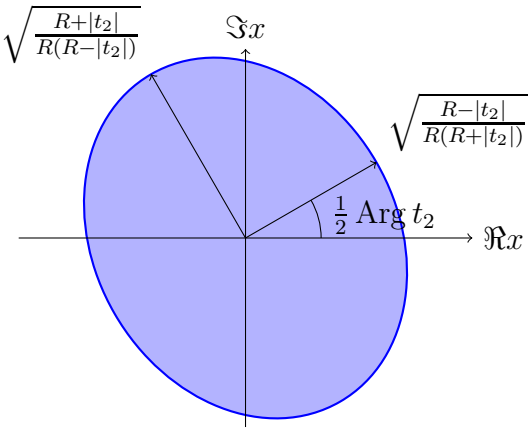
We assume that there are exactly two sheets, and that $S(x)$ is a solution of the algebraic equation $S^2 - S(S_+ + S_-) + S_+S_- = 0$. The coefficients of that equation must be single-valued, and analytic outside $\operatorname{supp}\bar{\rho}$. Assuming the minimal possible singularities, the coefficients must be polynomial functions of x. Together with their behaviour at ∞, this implies

$$S_+(x)S_-(x) = \frac{t_2}{\bar{t}_2}x^2 + \frac{t_2}{R^2} - \frac{1}{\bar{t}_2} \,, \tag{3.145}$$

$$S_+(x) + S_-(x) = -\left(\frac{t_2}{R} + \frac{R}{\bar{t}_2}\right)x \,. \tag{3.146}$$

Then the support of the equilibrium density is the ellipse with the equation

$$t_2 x^2 + \bar{t}_2 \bar{x}^2 + \left(\frac{t_2 \bar{t}_2}{R} + R \right) x \bar{x} = 1 - \frac{t_2 \bar{t}_2}{R^2} \ . \tag{3.147}$$



Polynomial potentials

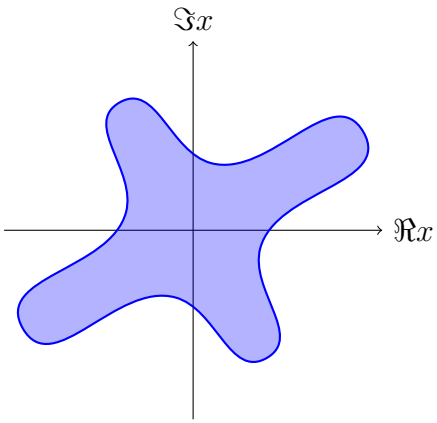
Let us assume that V and $\bar{V}$ are polynomials of degree $d+1$ with $V(x) = \sum_{k=1}^{d+1} \frac{t_k}{k} x^k$. There must be one sheet where $S(x)$ behaves as

$$S_1(x) \underset{x \to \infty}{=} -\frac{t_{d+1}}{R} x^d + \cdots \ . \tag{3.148}$$

On the other hand, using the analogous behaviour of $\bar{S}$, together with the equation (3.136), produces d sheets where

$$S_k(x) \underset{x \to \infty}{=} e^{\frac{2\pi i}{d} k} \left(-\frac{Rx}{\bar{t}_{d+1}} \right)^{\frac{1}{d}} + \cdots \ , \qquad k \in \{2, \dots, d+1\} \ . \tag{3.149}$$

We can look for a function $S(x)$ that lives on a $d+1$-sheeted cover of the complex plane, that is a solution of a polynomial equation of degree $d+1$. The corresponding support of the equilibrium density is the interior of an algebraic curve of degree $d+1$. For example, for $d = 4$, that support may look as follows:



In particular, this provides the solution of the Hele–Shaw problem.

Chapter 4

Loop equations

Loop equations are relations between correlation functions, that are obtained by integrating by parts in the matrix integral. Equivalently, loop equations follow from the invariance of the matrix integral under changes of integration variables – in this sense, loop equations are **Schwinger–Dyson equations**. Loop equations are useful because they provide an efficient way of computing correlation functions.

4.1 Exact loop equations

4.1.1 Loop equations

Consider the matrix integral

$$\mathcal{Z} = \int_\Gamma dM\, e^{-N \operatorname{Tr} V(M)} \ , \tag{4.1}$$

where N is the size of the matrix M. We assume that our matrix M belongs to a unitary ensemble ($\beta = 2$), see Section 4.1.5 for the cases $\beta \in \{1, 4\}$. We do not further specify the ensemble Γ over which the matrix M is integrated, or the potential $V(M)$, because loop equations are formally independent from these choices. As we will shortly see, we need only make sure that the integrand decreases fast enough at the boundaries of Γ.

We are interested in correlation functions of this model, such as the **moments**

$$\boxed{\left\langle \operatorname{Tr} M^k \right\rangle = \frac{1}{\mathcal{Z}} \int_\Gamma dM \left(\operatorname{Tr} M^k\right) e^{-N \operatorname{Tr} V(M)}} , \tag{4.2}$$

Assuming that the potential $V(M)$ is chosen so that the integrand vanishes at the boundaries of Γ, the integral of the following total derivative vanishes:

$$\sum_{i,j} \int_\Gamma dM \, \frac{\partial}{\partial M_{ij}} \left(\left(M^k\right)_{ij} e^{-N \operatorname{Tr} V(M)} \right) = 0 \ . \tag{4.3}$$

This apparently trivial equation actually leads to our first example of a loop equation. Computing the action of the derivative leads to

$$\sum_{i,j} \int_\Gamma dM \, \left(\sum_{l=0}^{k-1} \left(M^l\right)_{ii} \left(M^{k-l-1}\right)_{jj} - N \left(M^k\right)_{ij} \left(V'(M)\right)_{ji} \right) e^{-N \operatorname{Tr} V(M)} = 0 \ . \tag{4.4}$$

Rewriting this equation as a relation for correlation functions, we obtain the loop equation

$$\sum_{l=0}^{k-1} \Big\langle \operatorname{Tr} M^l \operatorname{Tr} M^{k-l-1} \Big\rangle - N \Big\langle \operatorname{Tr} M^k V'(M) \Big\rangle = 0 \ . \tag{4.5}$$

This equation can alternatively be obtained by performing the infinitesimal change of variable $M \to M + \epsilon M^k$ in the integral (4.1) at the order $O(\epsilon)$. The first term is the contribution of the Jacobian, and the second term comes from the variation of the integrand $e^{-N \operatorname{Tr} V(M)}$.

More general loop equations can be derived from the apparently trivial equation

$$\sum_{i,j} \int_\Gamma dM \, \frac{\partial}{\partial M_{ij}} \Big((M^{\mu_1})_{ij} \operatorname{Tr} M^{\mu_2} \cdots \operatorname{Tr} M^{\mu_n} \, e^{-N \operatorname{Tr} V(M)} \Big) = 0 \ . \tag{4.6}$$

Computing the action of the derivative, and rewriting the resulting equation in terms of correlation functions, we obtain the loop equation

$$\boxed{\begin{aligned} &\sum_{l=0}^{\mu_1 - 1} \Big\langle \operatorname{Tr} M^l \operatorname{Tr} M^{\mu_1 - l - 1} \prod_{i=2}^{n} \operatorname{Tr} M^{\mu_i} \Big\rangle \\ &\qquad + \sum_{j=2}^{n} \mu_j \Big\langle \operatorname{Tr} M^{\mu_1 + \mu_j - 1} \prod_{\substack{i=2 \\ i \neq j}}^{n} \operatorname{Tr} M^{\mu_i} \Big\rangle = N \Big\langle \operatorname{Tr} V'(M) M^{\mu_1} \prod_{i=2}^{n} \operatorname{Tr} M^{\mu_i} \Big\rangle \ . \end{aligned}} \tag{4.7}$$

The first term can be interpreted as splitting a trace $\operatorname{Tr} M^{\mu_1} \to \operatorname{Tr} M^l \times \operatorname{Tr} M^{\mu_1 - l - 1}$, and the second term as merging traces $\operatorname{Tr} M^{\mu_1} \times \operatorname{Tr} M^{\mu_j} \to \operatorname{Tr} M^{\mu_1 + \mu_j - 1}$. This equation can then be interpreted in terms of "splitting and merging" (or "cutting and joining") Feynman graphs. We will illustrate this graphical interpretation in Section 4.1.3. This is the reason why the Schwinger–Dyson equations for random matrices were given the particular name of loop equations by Migdal [49].

4.1.2 Space of solutions

Loop equations are linear equations for correlation functions, so their solutions form a vector space. In the case of a polynomial potential, we will show that this space is finite-dimensional, and compute its dimension.

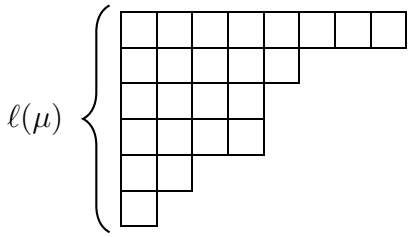
To do this, it is convenient to think of our matrix integrals in terms of the eigenvalues $\lambda_1, \cdots \lambda_N$ of the matrix M. The integrands are written as traces of functions of M, and are therefore symmetric polynomials of the eigenvalues. A single trace yields the polynomial

$$\operatorname{Tr} M^k = \sum_{i=1}^{N} \lambda_i^k \ , \tag{4.8}$$

and a product of traces yields the **power sum polynomial**

$$p_\mu(\lambda_1, \cdots \lambda_N) \stackrel{\text{def}}{=} \prod_{j=1}^{n} \operatorname{Tr} M^{\mu_j} = \prod_{j=1}^{n} \sum_{i=1}^{N} \lambda_i^{\mu_j} \ . \tag{4.9}$$

We label the power sum polynomial p_μ using the **partition** $\mu = (\mu_1, \ldots, \mu_n)$, where by definition $\mu_1 \geq \cdots \geq \mu_n$. The **length** of μ is $\ell(\mu) = n$, and its **weight** is $|\mu| = \sum_{i=1}^{n} \mu_i$. Partitions are graphically represented using **Young diagrams**:



In this example we represented the partition $\mu = (8, 5, 4, 4, 2, 1)$, with $\ell(\mu) = 6$, and $|\mu| = 24$.

Before taking loop equations into account, the space of correlation functions of our matrix model is linearly isomorphic to the space of the symmetric polynomials of N variables. A basis of this space is given by the power sum polynomials $\{p_\mu\}_{\ell(\mu)\leq N}$ – polynomials with lengths N or less linearly generate all symmetric polynomials because we have N variables $\lambda_1, \cdots \lambda_N$. The row lengths μ_i are however unbounded, so our space of correlation functions is infinite-dimensional.

Loop equations can now be interpreted as symmetric polynomials whose expectation values vanish, and they form a subspace of our space of symmetric polynomials. Finding a solution of the loop equations means giving an expectation value to all symmetric polynomials, such that loop equations have zero expectation value. In other words, solutions of the loop equations are one-forms whose kernels include the space of loop equations, and they belong to the coset space

$$\boxed{\{\text{solutions of the loop equations}\} = \frac{\{\text{symmetric polynomials}\}}{\{\text{loop equations}\}}}\,. \tag{4.10}$$

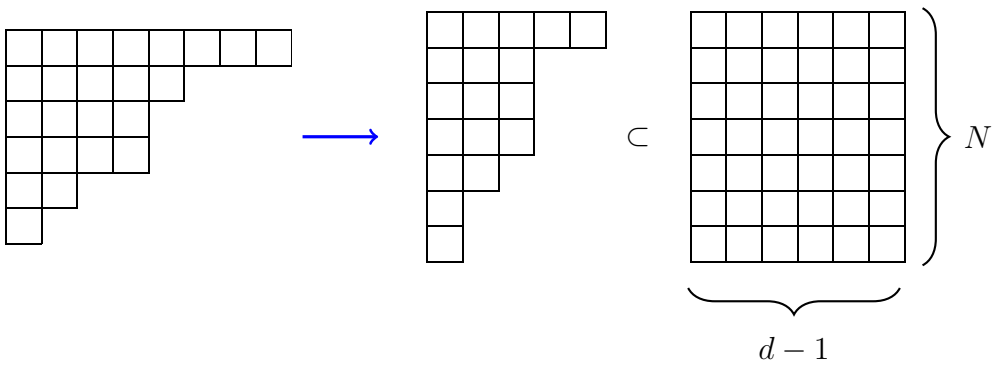
Let us now specialize to the case where the potential V is a polynomial of degree $d+1$,

$$V(M) = \sum_{k=1}^{d+1} \frac{t_k}{k} M^k\,. \tag{4.11}$$

(The constant term $k = 0$ is not written as it does not appear in loop equations.) Let us find a basis of the space of solutions. To do this, let us study how loop equations allow us to rewrite expectation values of power sum polynomials p_μ with $\ell(\mu) \leq N$ in terms of a subset of polynomials. Let us write the first loop equation (4.5) as

$$\sum_{l=0}^{\mu_1-1} \Big\langle \operatorname{Tr} M^l \operatorname{Tr} M^{\mu_1-l-1} \Big\rangle = N \Big\langle \operatorname{Tr} V'(M) M^{\mu_1} \Big\rangle = N \sum_{k=1}^{d+1} t_k \Big\langle \operatorname{Tr} M^{\mu_1+k-1} \Big\rangle\,. \tag{4.12}$$

This allows us to rewrite $\Big\langle \operatorname{Tr} M^{\mu_1+d} \Big\rangle$, which corresponds to a partition with one row of length $\mu_1 + d$, in terms of partitions with fewer boxes and shorter rows. The same argument, using the more general loop equations (4.7), shows that rows of lengths d or more can always be shortened, so that any polynomial is (up to loop equations) a linear combinations of power sum polynomials with fewer boxes and rows of lengths at most $d - 1$. In the process we may exceed our limit N for the number of rows. The point is that reducing the number of rows to N or less, using the fact that our polynomials have N variables, does not change the total degree of the polynomials – in other words, the number of boxes. Since shortening the rows also reduces the number of boxes, we can always reduce the numbers and lengths of rows until we obtain a linear combination of polynomials p_μ with $\mu_i \leq d - 1$ and $\ell(\mu) \leq N$. We illustrate this in the case $d = 7$ and $N = 7$:



So the polynomials p_μ where the partition μ fits in a box of size $N \times (d-1)$ are a spanning set of solutions of the loop equations. It can be shown that they are linearly independent modulo the loop equations, and therefore form a basis of solutions. The number of such partitions is the dimension of the space of solutions,

$$\boxed{d_N = \binom{N+d-1}{N}} \,. \tag{4.13}$$

This coincides with the dimension (1.45) of the space of convergent normal matrix integrals. So there is a one-to-one correspondence between solutions of loop equations, and integration contours Γ such that the matrix integral (4.1) converges.

4.1.3 Diagrammatic interpretation and Tutte's recursion

In this section we provide a diagrammatic interpretation for the loop equations. We consider a formal matrix integral (2.13), and rewrite its diagrammatic expression in terms of dual graphs – replacing a vertex of valence k with a k-gon:

$$\int_{\text{formal}} dM\, e^{-\frac{N}{2}\operatorname{Tr} M^2}\, e^{N\sum_{k=3}^{d+1}\frac{t_k}{k}\operatorname{Tr} M^k} \propto \sum_{G\in\mathcal{G}} \frac{N^{\chi(G)}}{\#\mathrm{Aut}(G)} \prod_{k=3}^{d+1} t_k^{n_k(G)} \,, \tag{4.14}$$

where $n_k(G)$ is the number of k-gons. Inserting the operator $\operatorname{Tr} M^{\mu_i}$ in the integral now amounts to inserting a marked μ_i-gon, and we have

$$\left\langle \frac{N}{\mu_1}\operatorname{Tr} M^{\mu_1} \cdots \frac{N}{\mu_n}\operatorname{Tr} M^{\mu_n} \right\rangle = \sum_{G\in\mathcal{G}_{\mu_1,\cdots\mu_n}} \frac{N^{\chi(G)}}{\#\,\mathrm{Aut}(G)} \prod_{k=3}^{d+1} t_k^{n_k(G)} \,. \tag{4.15}$$

Here $\mathcal{G}_{\mu_1,\cdots\mu_n}$ is the set of all discrete surfaces, connected or not, with n marked faces. The i-th marked face is supposed to be a μ_i-gon, with $\mu_i \geq 1$, and does not contribute to $n_{\mu_i}(G)$.

Let us now multiply eq. (4.15) with $N^{-n}\prod_i \mu_i$. The factor N^{-n} can be interpreted as replacing the Euler characteristic $\chi(G) = 2-2g$, where g is the genus, with the **modified Euler characteristic**, that is the characteristic of the graph that would be obtained by replacing the n marked faces with n holes,

$$\boxed{\chi'(G) = \chi(G) - n = 2 - 2g - n} \,. \tag{4.16}$$

Moreover, multiplying by μ_i amounts to counting the i-th marked face μ_i times, which is equivalent to choosing a marked edge. This greatly simplifies the combinatorics of graphs, as the presence of a single face with a marked edge ensures that our graph has no

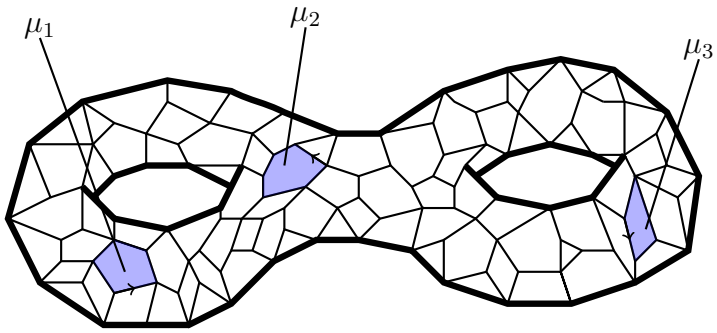
nontrivial automorphism. But we keep the factor $\frac{1}{\#\,\mathrm{Aut}(G)}$, because it can be nontrivial in the case $n = 0$ which we will later consider. Therefore we have

$$\left\langle \operatorname{Tr} M^{\mu_1} \cdots \operatorname{Tr} M^{\mu_n} \right\rangle = \sum_{G \in \hat{\mathcal{G}}_{\mu_1,\cdots\mu_n}} \frac{N^{\chi'(G)}}{\#\,\mathrm{Aut}(G)} \prod_{k=3}^{d+1} t_k^{n_k(G)}\,, \tag{4.17}$$

where $\hat{\mathcal{G}}_{\mu_1,\cdots\mu_n}$ is the set of discrete surfaces with n marked faces with marked edges. If we restricted the sum to the corresponding set $\mathcal{G}_{\mu_1,\cdots\mu_n}$ of *connected* discrete surfaces, we would compute a connected correlation function,

$$\boxed{\left\langle \operatorname{Tr} M^{\mu_1} \cdots \operatorname{Tr} M^{\mu_n} \right\rangle_{\mathrm{c}} = \sum_{G \in \mathcal{G}_{\mu_1,\cdots\mu_n}} \frac{N^{\chi'(G)}}{\#\,\mathrm{Aut}(G)} \prod_{k=3}^{d+1} t_k^{n_k(G)}}\,. \tag{4.18}$$

Let us draw an example of such a discrete surface, with genus $g = 2$ and $n = 3$ holes with $(\mu_1, \mu_2, \mu_3) = (5, 6, 4)$. Marked edges are denoted by arrows which follow the orientation of the surface, so that the hole is always on the left of the arrow:



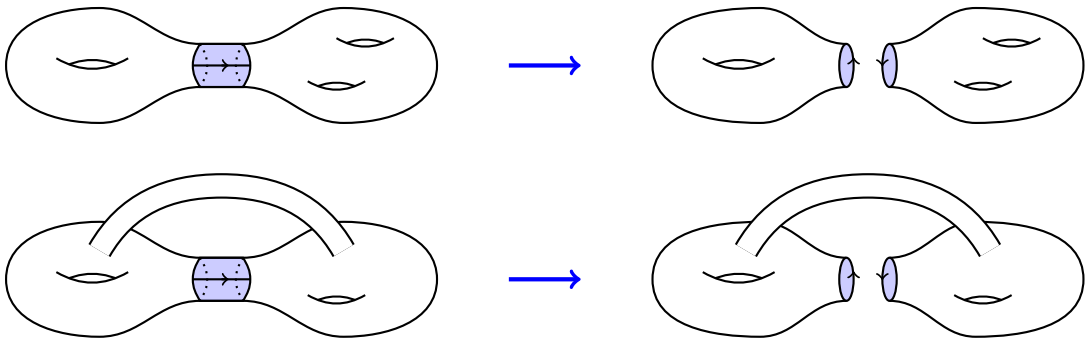
We insist that only the interior of a marked face is removed. For example, two neighbouring marked faces count as two holes, as the edge which separates them is not removed. Moreover, an edge between a marked face and itself is also not removed, and can suffice to keep our graph connected:

We will now study the diagrammatic interpretation of the loop equations in terms of discrete surfaces. Using the explicit form of the polynomial potential $V(M)$ which appears in our formal matrix integral (4.14), the loop equation (4.7) becomes

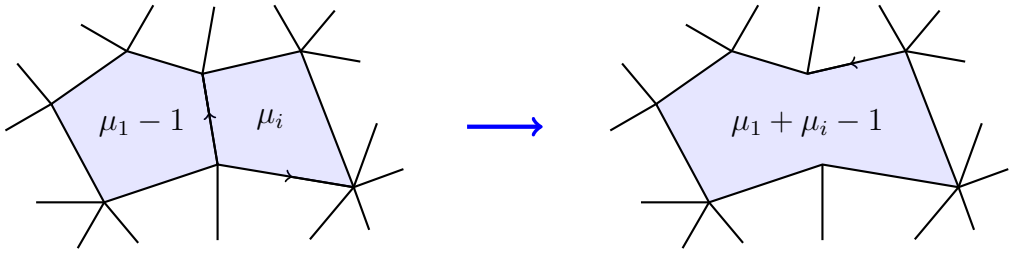
$$\begin{aligned} N\left\langle \operatorname{Tr} M^{\mu_1+1} \operatorname{Tr} M^{\mu_2} \cdots \operatorname{Tr} M^{\mu_n} \right\rangle &= \sum_{j=0}^{\mu_1-1} \left\langle \operatorname{Tr} M^{j} \operatorname{Tr} M^{\mu_1-j-1} \operatorname{Tr} M^{\mu_2} \cdots \operatorname{Tr} M^{\mu_n} \right\rangle \\ &+ \sum_{i=2}^{n} \mu_i \left\langle \operatorname{Tr} M^{\mu_1+\mu_i-1} \prod_{\substack{j=2 \\ j\neq i}}^{n} \operatorname{Tr} M^{\mu_j} \right\rangle + \sum_{k=3}^{d+1} t_k \left\langle \operatorname{Tr} M^{\mu_1+k-1} \operatorname{Tr} M^{\mu_2} \cdots \operatorname{Tr} M^{\mu_n} \right\rangle . \end{aligned} \tag{4.19}$$

This means that the discrete surfaces counted by the left-hand side are in bijection with discrete surfaces counted by the right-hand side. On the left-hand side, we have all discrete surfaces with n marked faces of respective lengths $\mu_1 + 1, \mu_2, \ldots, \mu_n$. On the right-hand side, we have three terms, which correspond to the three situations which can occur if we erase the marked edge of the first marked face:

1. If the erased marked edge is between the first marked face and itself, then we disconnect its boundary into two boundaries of lengths j and $\mu_1 - 1 - j$, and obtain discrete surfaces with $n+1$ boundaries. To get the correct combinatorics of surfaces, the two "daughter" marked faces must have well-defined marked edges, which we choose as the edges preceding and following the erased marked edge. So the first term of the right-hand side corresponds to *splitting* a marked face. We show two examples of this splitting procedure. In the first example, the surface gets disconnected. In the latter case, although it remains connected, its genus decreases by one.

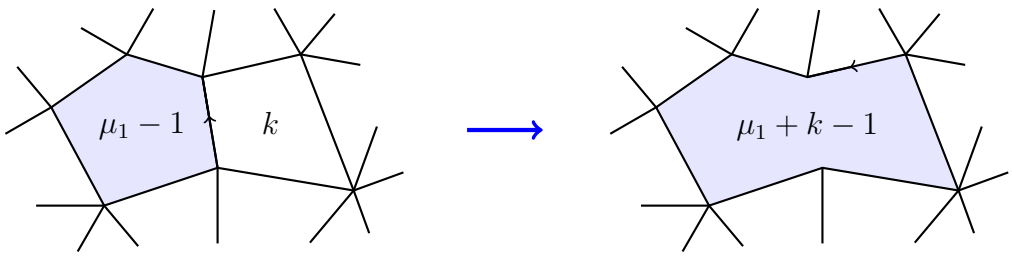


2. If the erased marked edge separates two different marked faces, then erasing it amounts to *merging* them, and we obtain a marked face with $\mu_1 + \mu_i - 1$ edges:



   We choose the edge following the erased edge as the marked edge of the new marked face. By doing that, we forgot the combinatorial information which was encoded in the marked edge of the i-th marked face. So we must restore a factor μ_i. This explains the second term of the right-hand side.

3. If there is an unmarked k-gon on the other side of the marked edge, then the number of marked faces does not change, but the first marked face becomes a $\mu_1 + k - 1$-gon:



   We choose the edge which follows the erased edge as our new marked edge. We must now include a factor t_k – the weight of the unmarked k-gon.

Applying this edge–erasing procedure recursively, all the discrete surfaces can be reduced to surfaces with very few edges. Counting discrete surfaces by recursively erasing edges is called **Tutte's recursion**.

4.1.4 Generating functions

The correlation functions discussed above are characterized by a set of integers $(\mu_1, \cdots \mu_n)$. It is often convenient to encode these correlation functions into analytic functions of variables $(x_1, \cdots x_n)$. For example, single-trace operators with $n = 1$ are encoded into the x-dependent operator

$$\operatorname{Tr} \frac{1}{x - M} = \sum_{\mu=0}^{\infty} x^{-\mu-1} \operatorname{Tr} M^{\mu} \ . \tag{4.20}$$

So we define the connected and disconnected n-point functions

$$\boxed{W_n(x_1, \cdots x_n) = \left\langle \operatorname{Tr} \frac{1}{x_1 - M} \cdots \operatorname{Tr} \frac{1}{x_n - M} \right\rangle_{\rm c} \ ,} \tag{4.21}$$

$$\boxed{\hat{W}_n(x_1, \cdots x_n) = \left\langle \operatorname{Tr} \frac{1}{x_1 - M} \cdots \operatorname{Tr} \frac{1}{x_n - M} \right\rangle \ .} \tag{4.22}$$

In particular we have $\hat{W}_1(x) = W_1(x)$, and

$$\hat{W}_2(x_1, x_2) = W_1(x_1) W_1(x_2) + W_2(x_1, x_2) \ . \tag{4.23}$$

Due to the appearance of $\operatorname{Tr} \frac{V'(x) - V'(M)}{x - M}$ in the loop equations, it is also useful to introduce

$$\boxed{P_n(x; x_1, \ldots, x_n) = \left\langle \operatorname{Tr} \frac{V'(x) - V'(M)}{x - M} \prod_{i=1}^{n} \operatorname{Tr} \frac{1}{x_i - M} \right\rangle_{\rm c} ,} \tag{4.24}$$

which is a polynomial in its first variable x.

Replacing the integer parameter k with a variable x in the first loop equation (4.5), we obtain

$$\left\langle \operatorname{Tr} \frac{1}{x - M} \operatorname{Tr} \frac{1}{x - M} \right\rangle - N \left\langle \frac{\operatorname{Tr} V'(M)}{x - M} \right\rangle = 0 \ . \tag{4.25}$$

In terms of the connected correlation functions W_n and P_n, this can be written as

$$W_2(x, x) + W_1(x)^2 = N \left(V'(x) W_1(x) - P_0(x) \right) \ . \tag{4.26}$$

Higher loop equations (4.7) can easily be written as linear equations for disconnected correlation functions $\hat{W}_n$,

$$\boxed{\hat{W}_{n+2}(x, x, I) + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \frac{\hat{W}_n(x, I - \{x_i\}) - \hat{W}_n(I)}{x - x_i} = N \Big(V'(x) \hat{W}_{n+1}(x, I) - \hat{P}_n(x; I) \Big) \ ,} \tag{4.27}$$

where $I = \{x_1, \cdots x_n\}$, and $\hat{P}_n$ is the disconnected version of the polynomial P_n (4.24). Rewriting this in terms of the connected correlation functions W_n yields the equations

$$\boxed{\begin{aligned} &W_{n+2}(x, x, I) + \sum_{J \subset I} W_{1+|J|}(x, J) \, W_{1+|I-J|}(x, I - J) \\ &\quad + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \frac{W_n(x, I - \{x_i\}) - W_n(I)}{x - x_i} = N \Big(V'(x) W_{n+1}(x, I) - P_n(x; I) \Big) \ . \end{aligned}} \tag{4.28}$$

Resolvents and eigenvalue densities

The operator $\mathrm{Tr}\,\frac{1}{x-M}$ is interesting not only because it encodes all the operators $\mathrm{Tr}\,M^k$, but also because it is related to the density of eigenvalues. Calling $\lambda_1, \cdots \lambda_N$ the eigenvalues of M, we indeed have

$$\mathrm{Tr}\,\frac{1}{x-M} = \sum_{i=1}^{N} \frac{1}{x-\lambda_i} = \int \frac{dx'}{x-x'} \sum_{i=1}^{N} \delta(x'-\lambda_i) = N \int \frac{dx'}{x-x'} \rho(x') \,. \tag{4.29}$$

So $\mathrm{Tr}\,\frac{1}{x-M}$ is (up to a factor N) the Stieltjes transform of the eigenvalue density $\rho(x)$. The expectation value $W_1(x) = \left\langle \mathrm{Tr}\,\frac{1}{x-M} \right\rangle$ is called the **resolvent** of the matrix model, and it contains all the relevant information about the spectrum of eigenvalues. In the saddle point approximation, we have $W_1(x) \sim N\omega(x)$ where the function $\omega(x)$ was introduced in eq. (3.8). If the saddle point approximation is valid, we can therefore use the results of Chapter 3 for computing the leading behaviour of $W_1(x)$ in the large N limit.

4.1.5 Arbitrary values of β, and the $\beta \to \frac{4}{\beta}$ duality

Let us generalize the loop equations to arbitrary values of the ensemble type β. We initially consider an integral over matrices M in an ensemble of type $\beta \in \{1, 2, 4\}$, where for convenience the potential comes with a factor β,

$$\mathcal{Z} = \int dM \; e^{-\frac{N\beta}{2} \mathrm{Tr}\, V(M)} \,. \tag{4.30}$$

When rewritten as an eigenvalue integral (1.35), this actually makes sense for all complex values of β such that the integral converges. The corresponding loop equations are β-dependent,

$$\begin{aligned} &\sum_{l=0}^{\mu_1-1} \Big\langle \mathrm{Tr}\, M^l \,\mathrm{Tr}\, M^{\mu_1-l-1} \prod_{i=2}^{n} \mathrm{Tr}\, M^{\mu_i} \Big\rangle + \left(\frac{2}{\beta} - 1\right) \mu_1 \Big\langle \mathrm{Tr}\, M^{\mu_1-1} \prod_{i=2}^{n} \mathrm{Tr}\, M^{\mu_i} \Big\rangle \\ &\quad + \frac{2}{\beta} \sum_{j=2}^{n} \mu_j \Big\langle \mathrm{Tr}\, M^{\mu_1+\mu_j-1} \prod_{\substack{i=2 \\ i\neq j}}^{n} \mathrm{Tr}\, M^{\mu_i} \Big\rangle = N \Big\langle \mathrm{Tr}\, V'(M) M^{\mu_1} \prod_{i=2}^{n} \mathrm{Tr}\, M^{\mu_i} \Big\rangle \,. \end{aligned} \tag{4.31}$$

Let us write these loop equations in terms of connected n-point functions. This is simpler if we include β-dependent normalization factors in the definitions of these n-point functions,

$$W_n(x_1, \ldots, x_n) = \left(\frac{\beta}{2}\right)^{\frac{n}{2}} \left\langle \prod_{i=1}^{n} \mathrm{Tr}\, \frac{1}{x_i - M} \right\rangle_c \,. \tag{4.32}$$

Then the β-dependent generalization of eq. (4.28) is

$$\begin{aligned} &W_{n+2}(x,x,I) + \sum_{J\subset I} W_{1+|I|}(x,J) W_{1+n-|J|}(x, I-J) + \left(\sqrt{\frac{\beta}{2}} - \sqrt{\frac{2}{\beta}}\right) \frac{\partial}{\partial x} W_{n+1}(x,I) \\ &\quad + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \frac{W_n(x, I-\{x_i\}) - W_n(I)}{x - x_i} = N\sqrt{\frac{\beta}{2}} \Big(V'(x) W_{n+1}(x,I) - P_n(x;I) \Big) \,. \end{aligned} \tag{4.33}$$

This only depends on N and β through two prefactors, which we now rewrite in terms of the Nekrasov variables (1.94),

$$\sqrt{\frac{\beta}{2}} - \sqrt{\frac{2}{\beta}} = \frac{\epsilon_1 + \epsilon_2}{\sqrt{-\epsilon_1 \epsilon_2}} \quad , \quad N\sqrt{\frac{\beta}{2}} = \frac{1}{\sqrt{-\epsilon_1 \epsilon_2}} \ . \tag{4.34}$$

These prefactors, and therefore the loop equations and their space of solutions, are invariant under the duality $\epsilon_1 \leftrightarrow \epsilon_2$. This supports the conjecture that the matrix integral is invariant under this duality.

4.2 Topological expansion of correlation functions

We will now discuss the topological expansions of matrix integrals and correlation functions. We will consider the two cases of formal and convergent matrix integrals. In each case we will study the behaviour of the corresponding correlation functions.

4.2.1 Formal integrals

In terms of connected correlation functions W_n, the diagrammatic expansion (4.18) becomes

$$W_n(x_1, \cdots x_n) = \frac{N\delta_{n,1}}{x_1} + \sum_{\mu_1, \cdots \mu_n = 1}^{\infty} \sum_{G \in \mathcal{G}_{\mu_1, \cdots \mu_n}} \frac{N^{\chi'(G)}}{\# \operatorname{Aut} G} \frac{\prod_{k=3}^{d+1} t_k^{n_k(G)}}{\prod_{i=1}^{n} x_i^{\mu_i + 1}} \ . \tag{4.35}$$

Here the sums run from $\mu_i = 1$ in contrast to eq. (4.20) where the sum run from $\mu = 0$, because $\mu_i = 0$ terms do not appear in connected correlation functions, except the term $\frac{N}{x_1}$ of W_1 which we wrote outside the sum. Partitioning our graphs according to the genus $\mathcal{G}_{\mu_1, \cdots \mu_n} = \sqcup_{g=0}^{\infty} \mathcal{G}_{g;\mu_1, \cdots \mu_n}$, we define $W_{g,n}$ as a sum over the set $\mathcal{G}_{g;\mu_1, \cdots \mu_n}$ of connected discrete graphs of genus g with n marked faces with marked edges,

$$W_{g,n}(x_1, \cdots x_n) = \frac{\delta_{g,0}\delta_{n,1}}{x_1} + \sum_{\mu_1, \cdots \mu_n = 1}^{\infty} \sum_{G \in \mathcal{G}_{g;\mu_1, \cdots \mu_n}} \frac{1}{\# \operatorname{Aut} G} \frac{\prod_{k=3}^{d+1} t_k^{n_k(G)}}{\prod_{i=1}^{n} x_i^{\mu_i + 1}} \ . \tag{4.36}$$

Equivalently, $W_{g,n}$ is the coefficient of $N^{\chi'(G)}$ in eq. (4.35), and we have the equality of formal series in $\mathbb{C}[[t_3, \ldots, t_{d+1}]]$,

$$\boxed{W_n = \sum_{g=0}^{\infty} N^{2-2g-n} W_{g,n}} \ , \tag{4.37}$$

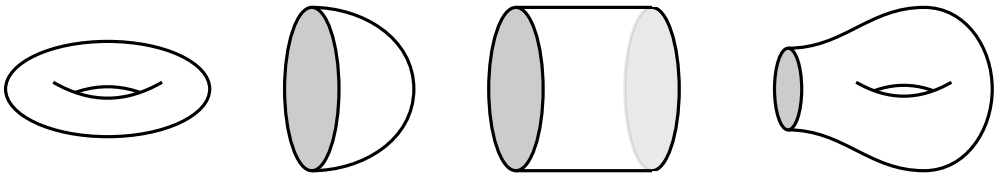
where the coefficient of each monomial $\prod_k t_k^{n_k}$ is a finite sum in both sides of the equality. (One cannot build discrete surfaces of arbitrarily high genus with a given number of faces and one cannot build infinitely many graphs with a given number of faces). In the case $n = 0$, this equality is valid provided we make the identification $W_0 = \log \mathcal{Z}$. Introducing moreover the notation $F_g = W_{g,0}$, we have

$$\boxed{\log \mathcal{Z} = \sum_{g} N^{2-2g} F_g} \ . \tag{4.38}$$

This equality of formal series in $\mathbb{C}[[t_3, \ldots, t_{d+1}]]$ should not be mistaken for a large N asymptotic expansion of a convergent matrix integral, see the discussion in Section 2.1.2.

Topology of the surfaces that contribute to $W_{g,n}$

If we consider marked faces as holes, the discrete surfaces that contribute to $W_{g,n}$ have genus g and n boundaries. We will use a sketch of such a surface as a diagrammatic representation of $W_{g,n}$. For example, we will use a torus for F_1, a disc for $W_{0,1}$, a cylinder for $W_{0,2}$, and a torus minus a disc for $W_{1,1}$:



F_1 $W_{0,1}$ $W_{0,2}$ $W_{1,1}$

Analytic structure of $W_{g,n}$

The correlation function $W_{g,n}$ is a formal series in $\mathbb{C}[[t_3, \dots, t_{d+1}]]$, and each one of its coefficients is a rational function of $x_1, \dots, x_n$, with poles only at $x_i = 0$. The residues at these poles is given by the first term in eq. (4.36), and therefore vanishes except in the case of $W_{0,1}$. Therefore, for any t_k-independent closed contour $\mathcal{C}$, we have

$$\frac{1}{2\pi i} \oint_{\mathcal{C}} W_{g,n}(x_1, \dots, x_n)\, dx_1 = \begin{cases} 1 & (g = 0,\ n = 1 \text{ and } \mathcal{C} \text{ surrounds } x_1 = 0) \\ 0 & (\text{otherwise}) \end{cases} . \tag{4.39}$$

Moreover, we have

$$\boxed{W_{g,n}(x_1, \dots, x_n) \underset{x_1 \to \infty}{=} \frac{\delta_{g,0}\delta_{n,1}}{x_1} + O\left(\frac{1}{x_i^2}\right)} . \tag{4.40}$$

This behaviour can actually be derived from the definition (4.21) of W_n, and is therefore valid not only for formal matrix integrals, but also for convergent matrix integrals.

4.2.2 Convergent matrix integrals

We now assume that the partition function $\mathcal{Z}$, and therefore also the correlation functions W_n, are no longer formal series, but convergent matrix integrals. Do they have expansions in powers of N? We will discuss this question, starting with the case of $\log \mathcal{Z}$. It can be proved that the leading term is always well-defined,

$$F_0 = \lim_{N \to \infty} \frac{1}{N^2} \log \mathcal{Z} . \tag{4.41}$$

The behaviour of the subleading terms depends on the potential V and on the integration contour Γ:

- In some cases, $\log \mathcal{Z}$ and W_n have large N asymptotic expansions (in other words, Poincaré expansions), and these expansions are formally given by the equations (4.37) and (4.38). In general, these asymptotic series are divergent, and typically $F_g \underset{g \to \infty}{\simeq} (2g)!$.

  It can be proved that such asymptotic expansions exist for the Hermitian ensemble $\Gamma = H_N$, provided the potential V is real and convex on the real axis. However, asymptotic expansions also exist in other cases, typically when the support of the equilibrium density is connected (one-cut case).

- In other cases, there is no such asymptotic expansion. Instead, the partition function can behave

$$\log \mathcal{Z} = N^2 F_0 + \left(F_1 + \sum_k \tilde{F}_{1,k} e^{-A_k N} \right) + o(1) \ , \tag{4.42}$$

  where $\Re A_k \geq 0$. In particular the case $\Re A_k = 0$ is possible, and we can therefore have oscillatory terms of the type of $\cos AN$ with $A \in \mathbb{R}$, which have no perturbative large N expansions. These oscillations typically appear when the support of the equilibrium density is disconnected (multi-cut case) [50]. We will study their properties in Section 4.5.

Partition functions with oscillatory terms can nevertheless be accessible to methods that generate asymptotic expansions. This is because the loop equations are linear constraints on the non-connected correlation functions $\hat{W}_n$. Therefore, all we need is a basis of d_N solutions, which do have perturbative expansions. Each solution corresponds to particular values of the polynomials P_n that appear in the loop equations, and choosing a basis amounts to choosing these polynomials. In particular, a basis of solutions can be constructed from formal integrals, as we will see in section 4.5 below. In this sense, solving loop equations for formal series is enough for solving them in general.

A general partition function, and in particular a convergent matrix integral, can therefore be written as a finite sum

$$\mathcal{Z} = \sum_{i=1}^{d_N} \mathcal{Z}^{(i)} \ , \tag{4.43}$$

where $\log \mathcal{Z}^{(i)}$ has a perturbative expansion of the type (4.38), $\log \mathcal{Z}^{(i)} = \sum_{g=0}^{\infty} N^{2-2g} F_g^{(i)}$. Assuming $\Re F_0^{(1)} = \max\{\Re F_0^{(i)}\}$, we then have

$$\log \mathcal{Z} = \log \mathcal{Z}^{(1)} + \log \left(1 + \sum_{i=2}^{d_N} \frac{\mathcal{Z}^{(i)}}{\mathcal{Z}^{(1)}} \right) \ , \tag{4.44}$$

$$= \sum_{g=0}^{\infty} N^{2-2g} F_g^{(1)} + \log \left(1 + \sum_{i=2}^{d_N} e^{-N^2 (F_0^{(1)} - F_0^{(i)})} e^{F_1^{(1)} - F_1^{(i)}} \left(1 + O\left(\frac{1}{N^2} \right) \right) \right) \ , \tag{4.45}$$

Since the summation over $i \in \{2, \dots, d_N\}$ is discrete, and the possible values of i depend on N, we typically have $F_0^{(1)} - F_0^{(i)} = O(\frac{1}{N})$, which leads to the non-perturbative terms for $\log \mathcal{Z}$, as in eq. (4.42). (See Section 4.5 for more details.)

In the cases of $W_{n \geq 1}$, the existence of a convergent perturbative expansion is determined by the existence of such an expansion for $\log \mathcal{Z}$. If oscillatory terms are present, we must write the matrix integral as a linear combination of integrals over other ensembles, and we find

$$\hat{W}_n(x_1, \cdots x_n) = \sum_i \frac{\mathcal{Z}^{(i)}}{\mathcal{Z}} \hat{W}_n^{(i)}(x_1, \cdots x_n) \ . \tag{4.46}$$

The correlation functions $\hat{W}_n^{(i)}$ have perturbative expansions, but the factors $\frac{\mathcal{Z}^{(i)}}{\mathcal{Z}}$ produce non-perturbative corrections.

4.3 Perturbative solution of the loop equations

Let us look for solutions of the loop equations that have topological expansions in powers of N. In order to find a basis of solutions, we only need to find enough independent solutions.

The loop equations (4.28) are not closed equations: to compute W_0 we need W_1, then W_2, and so on. However, inserting the expansion (4.37) of W_n, we will obtain closed equations for the coefficients $W_{g,n}$. Before we explain the general solution of these equations in the next Section, we will now solve them for a few low values of g and n.

To do this, we have to take care of the unknown polynomials P_n that appear in the loop equations. These polynomials have the topological expansions

$$\boxed{P_n = \sum_{g=0}^{\infty} N^{1-2g-n} P_{g,n}} \,, \tag{4.47}$$

where $P_{g,n}$ appears in the equation for $W_{g,n+1}$, and in particular $W_{0,1}$ depends on $P_{0,0}$. To obtain a basis of solutions, it is enough to allow $P_{0,0}$ to take arbitrary values in the space of polynomials of degree $\deg V''$ such that $P_{0,0}(x) \underset{x\to\infty}{\sim} \frac{V'(x)}{x}$. The remaining $P_{g,n}$ can then be chosen to our convenience. In particular, we can choose $P_{g,n}$ such that $W_{g,n}$ has the same singularities as $W_{0,1}$, and has vanishing integrals on contours $\mathcal{A}_i$ around the cuts of $W_{0,1}$,

$$\boxed{\forall\,(g,n) \neq (0,1), \qquad \oint_{\mathcal{A}_i} W_{g,n}(x_1,\dots,x_n)dx_1 = 0} \,. \tag{4.48}$$

(Compare with the property (4.39) of formal integrals.)

Allowing $P_{0,0}$ to take arbitrary values amounts to allowing the filling fractions

$$\boxed{\epsilon_i = \frac{1}{2\pi i}\oint_{\mathcal{A}_i} W_{0,1}(x)dx} \,, \tag{4.49}$$

to take arbitrary complex values such that $\sum_i \epsilon_i = 1$. What happens if we insist that $\epsilon_i \in \frac{1}{N}\mathbb{N}$ will be discussed in Section 4.5.

4.3.1 The disc amplitude $W_{0,1}$

We start with the first loop equation,

$$W_1(x)^2 + W_2(x,x) = N\Big(V'(x)W_1(x) - P_0(x)\Big) \,. \tag{4.50}$$

This is not a closed equation for $W_1(x)$, as it also involves $W_2(x,x)$. In the topological expansion, this equation yields infinitely many equations parametrized by $g \geq 0$, with the g-th equation being the coefficient of N^{2-2g},

$$\sum_{h_1+h_2=g} W_{h_1,1}(x)W_{h_2,1}(x) + W_{g-1,2}(x,x) = V'(x)W_{g,1}(x) - P_{g,0}(x) \,. \tag{4.51}$$

In the first equation, which is obtained for $g=0$, the term $W_{g-1,2}(x,x)$ is actually absent, and we obtain

$$W_{0,1}(x)^2 = V'(x)W_{0,1}(x) - P_{0,0}(x) \,. \tag{4.52}$$

This is a closed equation for $W_{0,1}$, for any choice of $P_{0,0}$. This equation is formally identical to the saddle point equation (3.31) for the resolvent $\bar{\omega}$, and we reproduce the solution

$$\boxed{W_{0,1}(x) = \frac{1}{2}\left(V'(x) - M(x)\sqrt{\sigma(x)}\right)}\,, \tag{4.53}$$

where the polynomials $M(x)$ and $\sigma(x)$ are such that $(V')^2 - 4P_{0,0} = M^2\sigma$, with $\sigma(x)$ having only simple roots.

In the case of formal integrals, the equation (4.39) for $W_{0,1}$ is incompatible with the existence of cuts that do not contain the point $x = 0$. We must therefore have only one cut, and that cut must contain the point $x = 0$. This uniquely determines $P_{0,0}(x)$ as a polynomial with coefficients in $\mathbb{C}[[t_3, \dots, t_{d+1}]]$.

4.3.2 The cylinder amplitude $W_{0,2}$

The leading term of the second loop equation is

$$2W_{0,2}(x,x')W_{0,1}(x) + \frac{\partial}{\partial x'}\frac{W_{0,1}(x) - W_{0,1}(x')}{x - x'} = V'(x)W_{0,2}(x,x') - P_{0,1}(x;x')\,. \tag{4.54}$$

Now that we know $W_{0,1}$, this is a closed equation for $W_{0,2}$, assuming we also know $P_{0,1}$. The solution is

$$W_{0,2}(x,x') = \frac{\frac{\partial}{\partial x'}\frac{W_{0,1}(x)-W_{0,1}(x')}{x-x'} + P_{0,1}(x;x')}{V'(x) - 2W_{0,1}(x)}\,, \tag{4.55}$$

$$= -\frac{1}{2(x-x')^2} + \frac{\frac{\partial}{\partial x'}\frac{M(x')\sqrt{\sigma(x')}}{2(x-x')} + \frac{\partial}{\partial x'}\frac{V'(x)-V'(x')}{2(x-x')} + P_{0,1}(x;x')}{M(x)\sqrt{\sigma(x)}}\,. \tag{4.56}$$

Since $P_{0,1}(x;x')$ is a polynomial of degree $d-1$ in x, $W_{0,2}(x,x')$ must be an algebraic function of x, and by symmetry also of x', and must live on the same Riemann surface as $W_{0,1}(x)$, which we will shortly call the spectral curve. For the purpose of finding a basis of solutions of the loop equations, we can furthermore choose $P_{0,1}$ to our will, subject to the constraint that the resulting $W_{0,2}(x,x')$ must be symmetric under the permutation of its two arguments. It is always possible to choose $P_{0,1}$ such that $W_{0,2}$ has not poles at the zeroes of $M(x)$, and is therefore singular only at the spectral edges – the zeros of $\sigma(x)$. This fully determines $P_{0,1}$ in the one-cut case. In the multi-cut case, that is whenever the genus of the spectral curve is strictly positive, additional conditions are needed in order to fully determine $P_{0,1}$ and therefore $W_{0,2}$. A convenient condition is eq. (4.48) for $W_{0,2}$.

Whatever the choice of $P_{0,1}$, from the solution (4.56) of the loop equation, we deduce that $W_{0,2}$ obeys the same Riemann–Hilbert equation (3.80) as the saddle point limit of the two-point function $\bar{\omega}_2$, and therefore must have an expression of the form of eq. (3.81).

4.3.3 Computing $W_{1,1}$

Once we know $W_{0,1}$ and $W_{0,2}$, we can find $W_{1,1}$ using the case $g = 1$ of the first loop equation (4.51),

$$2W_{0,1}(x)W_{1,1}(x) + W_{0,2}(x,x) = V'(x)W_{1,1}(x) - P_{1,0}(x)\,, \tag{4.57}$$

and the solution is

$$W_{1,1}(x) = \frac{W_{0,2}(x,x) + P_{1,0}(x)}{M(x)\sqrt{\sigma(x)}} \ . \tag{4.58}$$

This proves that $W_{1,1}(x)$ is an algebraic function, which again lives on the spectral curve, and whose sign flips when we change sheets. In order to determine it, we should choose $P_{1,0}$. From eq. (4.40), we deduce that $P_{1,0}$ is a polynomial of degree $d-2$. An assumption that fully determines $P_{1,0}$ is that $W_{1,1}$ has no simple poles at the zeros of $M(x)$, together with eq. (4.48) for $W_{1,1}$.

One-cut case

Let us consider the one-cut case. Let z be the Joukowsky variable, which is related to x via the Joukowsky map $x(z)$. The two spectral edges corresponds to $z = \pm 1$, and the involution that connects the two sheets corresponds to $z \to \frac{1}{z}$.

From Section 3.3.3, we have in the one-cut case

$$W_{0,2}(x,x')dxdx' = -B\left(z, \frac{1}{z'}\right) \ , \tag{4.59}$$

where $B(z,z') = \frac{dzdz'}{(z-z')^2}$ is the fundamental second kind differential of the spectral curve. We will now compute $W_{1,1}(x)dx$. Introducing a function $w(z)$ of the Joukowsky variable as

$$w(z) = -M(x(z))\sqrt{\sigma(x(z))} = W_{0,1}(x(z)) - W_{0,1}\left(x\left(\frac{1}{z}\right)\right) \ , \tag{4.60}$$

the solution (4.58) for $W_{1,1}(x)$ can be rewritten in terms of differential forms as

$$W_{1,1}(x)dx = \frac{W_{0,2}(x,x)dx^2 + P_{1,0}(x)dx^2}{-w(z)dx} \ . \tag{4.61}$$

In the Joukowsky variable, this differential form becomes

$$W_{1,1}(x)dx = \tilde{W}_{1,1}(z)dz = \frac{-B(z,\frac{1}{z}) + P_{1,0}(x(z))dx(z)^2}{-w(z)dx(z)} \ . \tag{4.62}$$

With our choice of $P_{1,0}$, this form can have poles only at $z = \pm 1$. Let us now rewrite it in terms of the fundamental second kind differential only, thereby eliminating the explicit dependence on $P_{1,0}$. Using the analyticity properties of $\tilde{W}_{1,1}(z)dz$, and its antisymmetry under $z \to \frac{1}{z}$, we compute

$$\tilde{W}_{1,1}(z) = \frac{1}{2\pi i}\oint_z \frac{dz'}{z'-z}\tilde{W}_{1,1}(z') = \sum_{\pm}\frac{1}{2\pi i}\oint_{\pm 1}\frac{dz'}{z-z'}\tilde{W}_{1,1}(z') \ , \tag{4.63}$$

$$= \frac{1}{2}\sum_{\pm}\frac{1}{2\pi i}\oint_{\pm 1}\left(\frac{1}{z-z'} - \frac{1}{z-\frac{1}{z'}}\right)\tilde{W}_{1,1}(z')dz' \ . \tag{4.64}$$

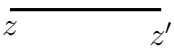
Inserting eq. (4.62) for $\tilde{W}_{1,1}(z')$, the contribution of $P_{1,0}$ now vanishes, because that polynomial has no poles at $z' = \pm 1$. Introducing moreover the **recursion kernel**

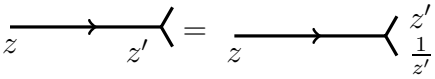
$$K(z,z') = \frac{1}{w(z')dx(z')}\int_{\frac{1}{z'}}^{z'} B(z,\cdot) = \frac{1}{w(z')dx(z')}\left(\frac{1}{z-z'} - \frac{1}{z-\frac{1}{z'}}\right)dz \ , \tag{4.65}$$

we finally find

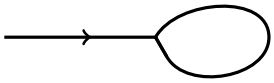
$$W_{1,1}(x)dx = \tilde{W}_{1,1}(z)dz = \frac{1}{2}\sum_{\pm}\frac{1}{2\pi i}\oint_{\pm 1} K(z,z')B\left(z',\frac{1}{z'}\right) \ . \tag{4.66}$$

This expression for $W_{1,1}(x)$ or equivalently of $\tilde{W}_{1,1}(z)$ in terms of the recursion kernel and the fundamental second kind differential can be generalized not only beyond the one-cut case, but also to all correlation functions $W_{g,n}$, as we will see in the next Section. Schematically,

$B(z,z') =$  (4.67)

$K(z,z') =$  (4.68)

where $B(z,z')$ is an unoriented propagator as it is invariant under the permutation of its arguments, while $K(z,z')$ is a kind of oriented propagator attached to a trivalent vertex with only two independent variables. Then

$W_{1,1} =$  (4.69)

4.4 Topological recursion

We will now introduce the topological recursion equation for the correlation functions $W_{g,n}$. Solving this equation leads to an expression for $W_{g,n}$ as a sum of graphs with g loops and n legs – thinned versions of surfaces with g handles and n boundaries. Each graph denotes an expression built from the fundamental second kind differential and the recursion kernel of the spectral curve. That spectral curve can be an arbitrary Riemann surface, in other words we can have arbitrarily many cuts.

The topological recursion originally arose in random matrix theory, but has much wider applications. For example, in topological string theory, correlation functions $W_{g,n}$ can be defined as integrals over moduli spaces of maps from Riemann surfaces of genus g with n punctures, to a Calabi–Yau manifold. These correlation functions can be computed using the topological recursion [18, 19], and they encode the Gromov–Witten invariants of our Calabi–Yau manifold. (The relation between random matrix theory and string theory had been noticed earlier by Dijkgraaf and Vafa [16].) Other invariants that appear in topological quantum field theories, such as the Reshetikhin–Turaev invariant and the Jones polynomial, are also expected to be computable using the topological recursion.

4.4.1 Spectral curve and differential forms

Spectral curve

The correlation function $W_{g,n}(x_1,\dots,x_n)$ is a multivalued function of x_i, which belong to $\mathbb{C}$ minus the cuts. It corresponds to a single-valued function of coordinates z_i on a Riemann surface Σ which is obtained by gluing several copies of $\mathbb{C}$ along the cuts, and which is called the spectral curve. (See Section 3.3.1.) This construction of Σ comes with a natural analytic map $x : \Sigma \to \mathbb{C}$ – in the one-cut case, the Joukowsky map.

The multivalued function $W_{0,1}(x)$ (4.53) gives rise to a single-valued, analytic map

$$\begin{array}{rccc} y: & \Sigma & \to & \mathbb{C} \\ & z & \mapsto & W_{0,1}(x(z)) \end{array} \tag{4.70}$$

and we now redefine the **spectral curve** as the triple (Σ, x, y). The maps x and y are related by the equation of the algebraic curve Σ,

$$\boxed{\left\{(x,y) \in \mathbb{C}\times\mathbb{C} \middle| y^2 - V'(x)y + P_{0,0}(x) = 0\right\} \quad = \quad \left\{(x(z), y(z)) \middle| z \in \Sigma\right\}} \,. \tag{4.71}$$

This equation is quadratic in y in our case of a one-matrix model, but topological recursion also works in more general cases.

Branch points

The analytic map x is locally invertible everywhere, except at the points where $x'(z) = 0$, that is the points where the differential $dx = x'(z)dz$ vanishes. These points are called the **branch points** of the spectral curve, and their images by the map x are the spectral edges.

In the generic case, the differential dx has a simple zero at a branch point a. If now z is a local complex coordinate on Σ, we have

$$x(z) = x(a) + \frac{x''(a)}{2}(z-a)^2 + \frac{x'''(a)}{6}(z-a)^3 + O\left((z-a)^4\right) \,. \tag{4.72}$$

Since x depends quadratically on z at the leading order, there must exist a nontrivial local map $\sigma_a(z)$ such that $x(z) = x(\sigma_a(z))$. This map is called the **local Galois involution** near the branch point a, and it reads

$$\sigma_a(z) = 2a - z - \frac{x'''(a)}{3x''(a)}(z-a)^2 + O\left((z-a)^3\right) \,. \tag{4.73}$$

This involution exchanges two sheets of Σ, in its construction as a multiple cover of the complex x-plane.

For example, in the one-cut case, derivating the Joukowsky map yields $dx = \gamma(1 - z^{-2})dz$. The two branch points are therefore $z = \pm 1$, and their images are indeed the spectral edges. The Joukowsky map obeys $x(z^{-1}) = x(z)$, and $z \mapsto z^{-1}$ is not just the local Galois involution near the branch points, but a global analytic map.

Differential forms

From the correlation function $W_{g,n}(x_1, \dots, x_n)$ and the map x, we build the following differential form on Σ,

$$\boxed{\omega_{g,n}(z_1, \cdots z_n) = \left(W_{g,n}(x_1, \cdots x_n) + \delta_{g,0}\delta_{n,2}\frac{1}{(x_1 - x_2)^2}\right) dx_1 \cdots dx_n} \,, \tag{4.74}$$

where

$$x_i = x(z_i) \quad \text{and} \quad dx_i = x'(z_i)dz_i \,. \tag{4.75}$$

The double pole that we add in the case of the two-point function $\omega_{0,2}$ allows it to coincide with the fundamental second kind differential of Σ,

$$\boxed{\omega_{0,2}(z_1, z_2) = B(z_1, z_2)} \,. \tag{4.76}$$

This was shown explicitly in the one-cut case in Section 3.3.3, and follows more generally from Section 4.3.2's results and assumptions on $W_{0,2}$. In particular, in contrast to the

forms $\omega_{g,n}$ with $(g,n) \neq (0,2)$, the form $\omega_{0,2}$ depends only on the Riemann surface Σ, and not on the maps x, y that also enter our definition of the spectral curve. Then, by definition of y, we have

$$\boxed{\omega_{0,1} = ydx}\ , \tag{4.77}$$

and the topological recursion will allow us to compute the rest of the forms $\omega_{g,n}$ from $\omega_{0,1}$ and $\omega_{0,2}$.

In the notations of algebraic geometry, we have (assuming $2g-2+n>0$)

$$\omega_{g,n} \in H^0(\Sigma^n, K_\Sigma(*\{a\})^{\boxtimes n})^{\mathfrak{S}_n}\ , \tag{4.78}$$

where $K_\Sigma(*\{a\})$ is the canonical bundle of meromorphic one-forms on Σ having arbitrary poles at the branch points $a \in \{a\}$, and the presence of the symmetric group $\mathfrak{S}_n$ indicates that $\omega_{g,n}(z_1, \cdots z_n)$ is symmetric under permutations of its n arguments.

4.4.2 Recursion relation

In terms of the differential forms $\omega_{g,n}$, the loop equations (4.28) become

$$\begin{aligned}\omega_{g-1,n+1}(z,\sigma_a(z),z_2,\cdots z_n) + \sum_{\substack{h+h'=g\\ I\sqcup I'=\{z_2,\cdots z_n\}}} \omega_{h,1+|I|}(z,I)\omega_{h',1+|I'|}(\sigma_a(z),I') \\ = dx(z)^2 P_{g,n}(x(z);x(z_2),\dots,x(z_n))dx(z_2)\dots dx(z_n) \\ - dx(z)^2 \sum_{j=2}^n dx(z_j)\frac{d}{dx(z_j)}\frac{\omega_{g,n-1}(z_2,\dots,z_n)}{(x(z)-x(z_j))dx(z_j)}\ .\end{aligned} \tag{4.79}$$

As a function of z, the right-hand side has no poles, and even vanishes at the spectral edges $z \to a$ due to the prefactor $dx(z)^2$. We choose $P_{g,n}$ so that $\omega_{g,n}(z,z_2,\dots,z_n)$ has no pole at the zeros of $M(x(z))$, and so that the $\mathcal{A}$-cycle integrals of $\omega_{g,n}$ vanish. Then $\omega_{g,n}$ has poles only at the spectral edges. Using similar calculations as we did for $\omega_{1,1}$ Section 4.3.3, we can eliminate $P_{g,n}$ and obtain the recursion relation

$$\boxed{\begin{aligned}\omega_{g,n}(z_1,\cdots z_n) = \sum_{a\in\{a\}} \operatorname*{Res}_{z=a} K_a(z_1,z)\Bigg[\omega_{g-1,n+1}(z,\sigma_a(z),z_2,\cdots z_n) \\ + {\sum_{\substack{h+h'=g\\ I\sqcup I'=\{z_2,\cdots z_n\}}}}' \omega_{h,1+|I|}(z,I)\omega_{h',1+|I'|}(\sigma_a(z),I')\Bigg]\ ,\end{aligned}} \tag{4.80}$$

where the modified summation $\sum'$ excludes the terms $(h,I) = (0,\emptyset)$ and $(h,I) = (g,\{z_2,\cdots z_n\})$, and σ_a is the local Galois involution near the branch point a. The recursion kernel, which we already encountered in the one-cut case (4.65), has the general expression

$$\boxed{K_a(z_1,z) = \frac{1}{2}\frac{\int_{\sigma_a(z)}^z \omega_{0,2}(z_1,\cdot)}{\omega_{0,1}(z)-\omega_{0,1}(\sigma_a(z))}}\ . \tag{4.81}$$

The recursion relation (4.80) is called **topological recursion** equation. It is called topological because it expresses a correlation function $\omega_{g,n}$ of modified Euler characteristic

$|\chi'| = 2g-2+n$ in terms of correlation functions of smaller characteristics $|\chi'| \leq 2g-3+n$. The recursion therefore stops after $2g-2+n$ steps, when the only appearing correlation function is $\omega_{0,2}$. Moreover, it can be proved that the recursion always produces correlation functions that are symmetric under permutations, although this is not manifest as the variable z_1 seems to play a special role in the recursion relation.

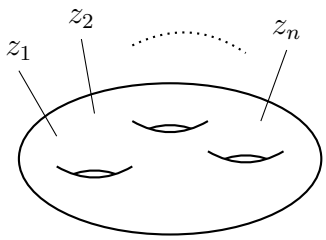
The free energies $F_g = \omega_{g,0}$ cannot be computed using the recursion formula. However, if $g \geq 2$ they are given in terms of $\omega_{g,1}$ by [51]

$$\boxed{F_g = \frac{1}{2-2g} \sum_a \operatorname*{Res}_{z=a} \omega_{g,1}(z) \int^z \omega_{0,1}(\cdot)} \,, \tag{4.82}$$

where the function $\int^z \omega_{0,1}(\cdot)$ is a primitive of $\omega_{0,1}$, and is supposed to be defined locally near each branch point. The result is independent of a choice of integration constant for the primitive, i.e. of the initial point of integration. This formula does not give F_0 and F_1. Particular expressions for these quantities can be found in [51].

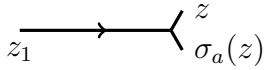
4.4.3 Graphical representation

We represent the genus g term of the n-point as a genus g Riemann surface with n marked points:

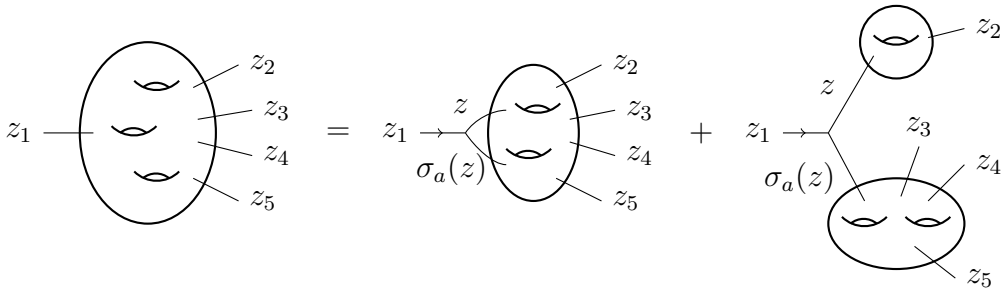
$\omega_{g,n}(z_1, \cdots z_n) =$ 

(g handles)

The fundamental second kind differential is still represented as an unoriented line (4.67). Generalizing the one-cut case (4.68), the recursion kernel is represented as

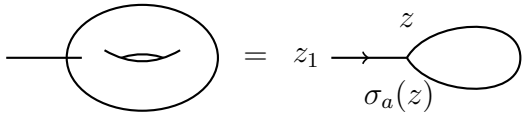
$K_a(z_1, z) \;=$ 

Then the recursion relation (4.80) is



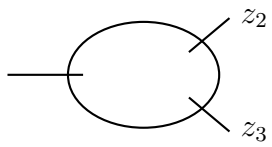
In this representation, the summations over the distributions of handles and marked points in the last term, are kept implicit. Also, as usual in the graphical representation of Feynman graphs, the vertex means an integration, here a sum of residues at all branch points.

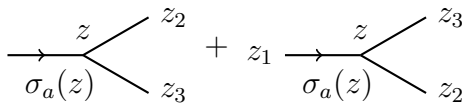
Let us show examples of computations of correlation functions, starting with the one-point function with $g = 1$:

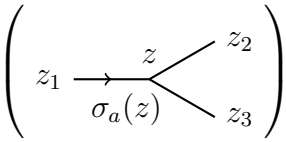
$\omega_{1,1}(z_1) =$  (4.83)

$$= \sum_a \operatorname*{Res}_{z=a} K_a(z_1, z) B(z, \sigma_a(z)) \ . \tag{4.84}$$

Our next example is the three-point function with $g = 0$:

$\omega_{0,3}(z_1, z_2, z_3) =$  (4.85)

$=$  (4.86)

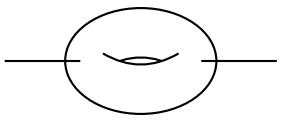
$= 2 \times$  (4.87)

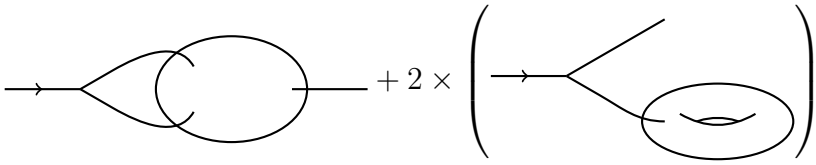
$$= 2 \sum_a \operatorname*{Res}_{z=a} K_a(z_1, z) B(z, z_2) B(\sigma_a(z), z_3) \ , \tag{4.88}$$

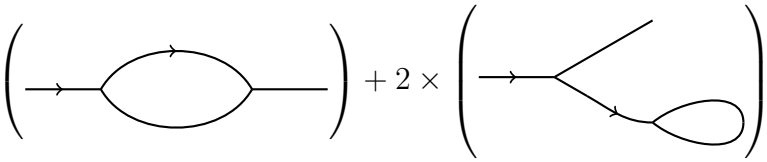
where the symmetry factor of 2 is due to the relation

$$K_a(z_1, z) = K_a(z_1, \sigma_a(z)) \ . \tag{4.89}$$

Our last example is the two-point function with $g = 1$. This makes use of the first two examples:

$\omega_{1,2} =$  (4.90)

$=$  (4.91)

$= 2 \times$  (4.92)

$$= 2 \sum_{a,b} \operatorname*{Res}_{z=a} \operatorname*{Res}_{z'=b} K_a(z_1, z) K_b(z, z') \times \Big(B(\sigma_a(z), \sigma_a(z')) B(z', z_2) + B(\sigma_a(z), z_2) B(z', \sigma_a(z')) \Big) \ . \tag{4.93}$$

Again, the relation (4.89) leads to symmetry factors of 2. The order matters in computing the residues at vertices, and the graphical rule is to compute the residues at vertices following the arrows backwards from leaves to root.

4.5 Non-perturbative asymptotics and oscillations

The topological recursion provides one solution of the loop equations for any choice of filling fractions $\epsilon_i \in \mathbb{C}$ with $\sum_i \epsilon_i = 1$, and the resulting solutions are linearly independent. However, in a matrix model of size N, the filling fractions are of the type $\epsilon_i = \frac{n_i}{N}$, where n_i is the number of eigenvalues inside the cycle $\mathcal{A}_i$, and $\#\{i|n_i \neq 0\}$ is the number of cuts. The number of choices of n_i is precisely the dimension d_N (4.13) of the space of solutions of the loop equations for the matrix model. Therefore, the topological recursion provides a basis of that space.

The correlation functions $W_{g,n}$ are smooth functions of the filling fractions ϵ_i, and are considered N-independent in the topological expansion, which is a perturbative expansion in powers of N. However, taking $\epsilon_i = \frac{n_i}{N}$ introduces an N-dependence in $W_{g,n}$, and therefore complicates the asymptotic behaviour of the partition function and correlation functions W_n. We will now see that summing over n_i gives rise to non-perturbative, oscillatory terms. We will find that these non-perturbative terms, like the perturbative terms, have universal properties.

4.5.1 Sum over filling fractions

Let us consider the path

$$\gamma = \sum_{i=1}^{s} c_i \gamma_i \quad \text{with} \quad c_i = e^{2\pi i \mu_i} \ , \tag{4.94}$$

where $\gamma_1, \dots, \gamma_s$ are independent contours. The normal matrix integral $\mathcal{Z}(\gamma^N)$ is a linear combination (1.46) of integrals $\mathcal{Z}(\vec{\gamma}^{\vec{n}})$ with fixed filling fractions $\epsilon_i = \frac{n_i}{N}$. Redefining the partition functions to eliminate factorial normalization factors, we write this combination as

$$\boxed{\mathcal{Z}(\gamma^N) = \sum_{n_1+\cdots+n_s=N} e^{2\pi i \sum_{i=1}^s n_i \mu_i} \mathcal{Z}(\vec{\gamma}^{\vec{n}})} \ . \tag{4.95}$$

We know that the logarithm of an integral $\mathcal{Z}(\vec{\gamma}^{\vec{n}})$ with fixed filling fractions has a topological expansion, so that

$$\boxed{\mathcal{Z}(\vec{\gamma}^{\vec{n}}) = \exp \sum_{g=0}^{\infty} N^{2-2g} F_g(\vec{\epsilon})} \ . \tag{4.96}$$

We now want to insert this expression in eq. (4.95), and deduce the expansion of $\log \mathcal{Z}(\gamma^N)$, including non-perturbative contributions.

As follows from its expression (4.82), $F_g(\vec{\epsilon})$ is an analytic function of its parameters, and in particular of the filling fractions $\vec{\epsilon}$. We will replace $F_g(\vec{\epsilon})$ with its Taylor expansion near a reference point $\vec{\epsilon}^*$, which we write as

$$F_g(\vec{\epsilon}) = F_g(\vec{\epsilon}^*) + \sum_{m=1}^{\infty} \frac{1}{m!} (\vec{\epsilon} - \vec{\epsilon}^*)^{\vec{m}} F_g^{(\vec{m})}(\vec{\epsilon}^*) \ , \tag{4.97}$$

$$= F_g(\vec{\epsilon}^*) + \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{1 \leq i_1 \cdots i_m \leq s} (\epsilon_{i_1} - \epsilon_{i_1}^*) \cdots (\epsilon_{i_m} - \epsilon_{i_m}^*) \, \partial_{i_1} \cdots \partial_{i_m} F_g(\vec{\epsilon}^*) \ . \tag{4.98}$$

In this formula the first line is a condensed notation, where in particular the sum over $i_1 \cdots i_m$ is kept implicit.

Since we want to sum over the numbers n_i of eigenvalues, we must rewrite the filling fractions in terms of n_i,

$$\epsilon_i - \epsilon_i^* = \frac{n_i - N\epsilon_i^*}{N} \ . \tag{4.99}$$

This suggests that we now consider $\epsilon_i = O(\frac{1}{N})$ instead of $\epsilon_i = O(1)$. Then the term $F_g^{(\vec{m})}(\vec{\epsilon})$ of $\log \mathcal{Z}(\vec{\gamma}^{\vec{n}})$ comes with a coefficient N^{2-2g-m} instead of N^{2-2g}. We will now only keep the terms with positive powers of N in the exponential (4.96), and use the Taylor expansion of the exponential function for all terms with negative powers of N. This leads to

$$\mathcal{Z}(\gamma^N) = e^{\sum_{g=0}^\infty N^{2-2g} F_g(\vec{\epsilon}^*)} \sum_{n_1+\cdots+n_s=N} e^{2\pi i \sum_{i=1}^s n_i \mu_i} \, e^{N(\vec{n}-N\vec{\epsilon}^*)F_0^{(\vec{1})}(\epsilon^*)} \, e^{\frac{1}{2}(\vec{n}-N\vec{\epsilon}^*)^2 F_0^{(\vec{2})}(\epsilon^*)}$$
$$\times \left(1 + \sum_{k=1}^\infty \frac{1}{k!} \sum_{\substack{m_i>0,\, g_i \\ 2g_i+m_i-2>0}} \prod_{i=1}^k \frac{N^{2-2g_i-m_i}}{m_i!} \prod_{i=1}^k (\vec{n}-N\vec{\epsilon}^*)^{\vec{m}_i} F_{g_i}^{(\vec{m}_i)}(\epsilon^*) \right) \ . \tag{4.100}$$

In each term of this expression, the dependence on the eigenvalue numbers $n_1, \ldots, n_s$ is a Gaussian function, times a polynomial. Therefore, the sum over $n_1, \ldots, n_s$ will produce theta functions.

Theta functions

Let τ be a $g \times g$ symmetric complex matrix whose imaginary part is positive definite:

$$\tau^T = \tau \quad , \quad \Im\tau > 0 \ . \tag{4.101}$$

The associated **theta function** is an entire function of a complex vector $\vec{u}$ of dimension g, which is defined by the absolutely convergent series

$$\boxed{\theta(\vec{u};\tau) = \sum_{\vec{n}\in\mathbb{Z}^g} e^{i\pi(\vec{n},\tau\vec{n})} \, e^{2\pi i(\vec{n},\vec{u})}} \ . \tag{4.102}$$

The theta function is even, $\mathbb{Z}^g$-periodic, and quasi-periodic in the direction $\tau\mathbb{Z}^g$,

$$\theta(-\vec{u}) = \theta(\vec{u}) \quad , \quad \theta(\vec{u}+\vec{n}+\tau\vec{m}) = e^{2\pi i(\vec{m},\vec{u})+i\pi(\vec{m},\tau\vec{m})}\theta(\vec{u}) \ . \tag{4.103}$$

Let us now write $\mathcal{Z}(\gamma^N)$ in terms of theta functions, starting from eq. (4.100). The relevant symmetric matrix and complex vector are

$$\boxed{\tau = \frac{1}{2\pi i} F_0^{(\vec{2})}(\vec{\epsilon}^*) \quad , \quad \vec{u}_0 = \vec{\mu} - \frac{N}{2\pi i} F_0^{(\vec{1})}(\vec{\epsilon}^*)} \ , \tag{4.104}$$

and in order to comply with eq. (4.101) we assume that $\Re F_0^{(\vec{2})}(\vec{\epsilon}^*)$ is negative definite, in other words that $-\Re F_0(\vec{\epsilon})$ is locally convex near $\vec{\epsilon} = \vec{\epsilon}^*$. The sum in eq. (4.100) is only on finitely many values of $\vec{n}$, but the values that are not of the type $N\epsilon^* + O(1)$ give exponentially small contributions. By adding such contributions, we can extend the sum to $\vec{n} \in \mathbb{Z}^{s-1}$. The leading term of $\mathcal{Z}(\gamma^N)$ is then approximated by the theta function

$$\mathcal{Z}(\gamma^N)\big|_{\text{leading}} \sim e^{N^2 F_0(\vec{\epsilon}^*)+F_1(\vec{\epsilon}^*)} e^{\frac{N^2}{2}(\vec{\epsilon}^*, F_0^{(\vec{2})}(\vec{\epsilon}^*)\vec{\epsilon}^*)} e^{-N^2 \vec{\epsilon}^* F_0^{(\vec{1})}(\vec{\epsilon}^*)} \theta\left(\vec{u}_0;\tau\right) \ . \tag{4.105}$$

The periodicity properties of the theta function imply that $\mathcal{Z}(\gamma^N)$ is quasi-periodic as a function of N, with oscillations of frequency $\frac{F_0^{(\vec{1})}(\vec{\epsilon}^*)}{2\pi i}$ if that vector is real. (If it is not real, we have exponential behaviour instead.)

By the same method, the subleading terms of $\mathcal{Z}(\gamma^N)$ can be approximated by derivatives of the theta function, and we obtain

$$\boxed{\begin{aligned}\mathcal{Z}_N(\gamma) \sim\ & e^{\sum_{g=0}^{\infty} N^{2-2g} F_g(\vec{\epsilon}^*)} e^{\pi i N^2(\vec{\epsilon}^*, \tau\vec{\epsilon}^*)} e^{2\pi i N(\vec{\epsilon}^*, (2\vec{u}_0 - \vec{\mu}))} \\ & \times \left(1 + \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{\substack{m_i>0, g_i \\ 2g_i+m_i-2>0}} \prod_{i=1}^{k} \frac{N^{2-2g_i-m_i}}{m_i!} F_{g_i}^{(\vec{m}_i)}(\vec{\epsilon}^*) \partial_{\vec{u}}^{\vec{m}_i}\right) \Theta(\vec{u};\tau)\Big|_{\vec{u}=2\pi i\vec{u}_0}\end{aligned}}, \tag{4.106}$$

where we introduced the twisted theta function

$$\boxed{\Theta(\vec{u};\tau) = e^{-N(\vec{\epsilon}^*, \vec{u})} \theta\left(\frac{1}{2\pi i}\vec{u};\tau\right)}. \tag{4.107}$$

Schematically, we can write the first few terms as

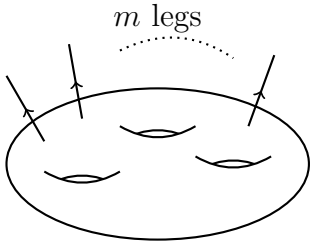
$$\begin{aligned}\mathcal{Z}_N(\gamma) \sim\ & e^{\sum_{g=0}^{\infty} N^{2-2g} F_g} e^{\pi i N^2(\vec{\epsilon}^*, \tau\vec{\epsilon}^*)} e^{2\pi i N(\vec{\epsilon}^*, (2\vec{u}_0 - \vec{\mu}))} \\ & \times \left(\Theta + \frac{1}{N}\left(F_1^{(\vec{1})}\Theta^{(\vec{1})} + \frac{1}{6}F_0^{(\vec{3})}\Theta^{(\vec{3})}\right) + O\left(\frac{1}{N^2}\right)\right).\end{aligned} \tag{4.108}$$

4.5.2 Graphical representation

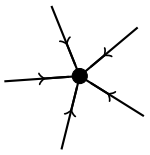
Partition function

The expansion (4.106) of the partition function has an efficient graphical representation, where

- $F_g^{(\vec{m})}$ corresponds to a surface of genus g with m outgoing legs, which we call a vertex,

$F_g^{(\vec{m})} =$ 

- $\Theta^{(\vec{m})}$ corresponds to a black dot with m incoming legs,

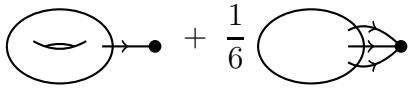
$\Theta^{(\vec{m})} =$  m legs

The partition function can then be written as a sum over all possible graphs with one black dot and arbitrarily many vertices,

$$\boxed{\begin{aligned}\mathcal{Z}_N(\gamma) \sim\ & e^{\sum_{g=0}^{\infty} N^{2-2g} F_g} e^{\pi i N^2(\vec{\epsilon}^*, \tau\vec{\epsilon}^*)} e^{2\pi i N(\vec{\epsilon}^*, (2\vec{u}_0 - \vec{\mu}))} \\ & \times \left(\Theta + \sum_{\mathcal{G}} \frac{N^{\chi(\mathcal{G})}}{\#\mathrm{Aut}(\mathcal{G})} \left[\prod_{\text{vertices } v} F_{g_v}^{(\vec{m}_v)}\right] \Theta^{(\sum_v \vec{m}_v)}\right),\end{aligned}} \tag{4.109}$$

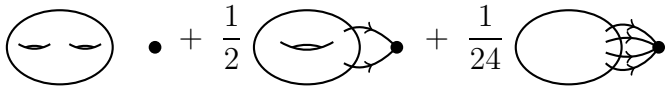
where $\mathrm{Aut}(\mathcal{G})$ is the group of automorphisms of the graph $\mathcal{G}$, and $\chi(\mathcal{G}) = \sum_v (2-2g_v-m_v)$ is its Euler characteristic. For example, the first two orders are:

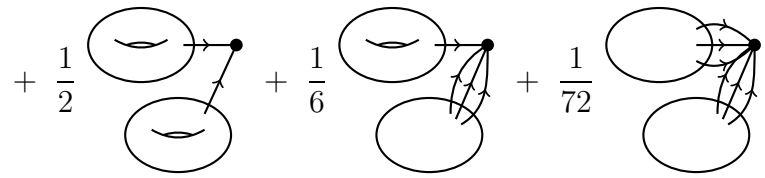
$O(N^{-1})$:



(4.110)

$O(N^{-2})$:





(4.111)

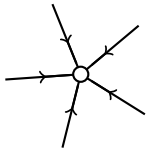
Free energy

While the partition function $\mathcal{Z}_N(\gamma)$ is linear in the function Θ and its derivatives, the free energy $\log \mathcal{Z}_N(\gamma)$ involves products of arbitrarily many theta functions. Graphically, this corresponds to replacing the unique black dot with an arbitrary number of white dots. More precisely, we define the derivative of $\log \Theta$,

$$T^{(\vec{m})}(\vec{u}) = \partial_{\vec{u}}^{\vec{m}} \log \Theta(\vec{u}) \ , \tag{4.112}$$

which we represent as a white dot with m ingoing legs:

$T^{(\vec{m})}$ =

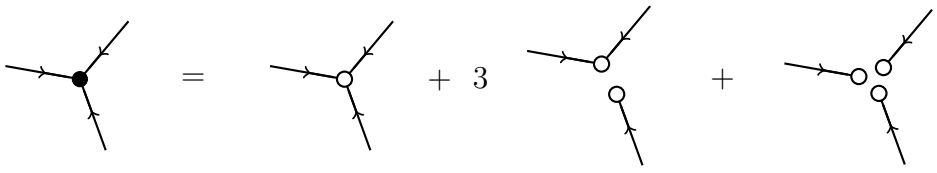


m legs

The relation between the derivatives of Θ and $\log \Theta$,

$$\frac{\Theta^{(\vec{m})}}{\Theta} = \sum_{\sqcup \vec{m}_i = \vec{m}} \prod_i T^{(\vec{m}_i)} \ , \tag{4.113}$$

amounts to decomposing the black dot into white dots, for instance,

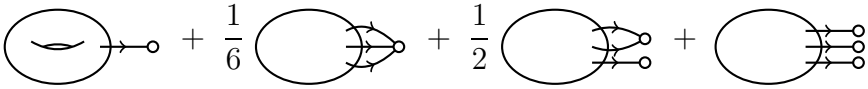


The free energy can then be written as a sum over all connected graphs with arbitrarily many white dots and vertices,

$$\log \mathcal{Z}_N(\gamma) \sim \sum_{g=0}^{\infty} N^{2-2g} F_g + \pi i N^2(\vec{\epsilon}^*, \tau\vec{\epsilon}^*) + 2\pi i N(\vec{\epsilon}^*, (2\vec{u}_0 - \vec{\mu}))$$
$$+ \log \Theta + \sum_{\mathcal{G}\,\text{connected}} \frac{N^{\chi(\mathcal{G})}}{\#\mathrm{Aut}(\mathcal{G})} \prod_{\text{vertices } v} F_{g_v}^{(\vec{m}_v)} \prod_{\text{white dots } v'} T^{(\vec{m}_{v'})} \ . \tag{4.114}$$

For example, the leading nontrivial contribution is a sum over the following graphs:

$O(N^{-1})$:



(4.115)

4.5.3 Universality and background independence

Using eq. (4.82), it is possible not only to prove that F_g is an analytic function of the filling fractions $\vec{\epsilon}$, but also to derive an expression for its derivatives as $\mathcal{B}$-cycle integrals,

$$\boxed{\frac{\partial F_g}{\partial \epsilon_i} = \oint_{\mathcal{B}_i} \omega_{g,1}} \,, \tag{4.116}$$

and more generally,

$$\frac{\partial^k F_g}{\partial \epsilon_{i_1} \dots \partial \epsilon_{i_k}} = \oint_{\mathcal{B}_{i_1}} \dots \oint_{\mathcal{B}_{i_k}} \omega_{g,k} \,. \tag{4.117}$$

Therefore, all perturbative contributions to the free energy have a universal expression in terms of the complex structure of the spectral curve. Then the non-perturbative oscillatory terms are universal too, since they are computed by summing the perturbative contributions over filling fractions.

Finally, let us come back to the choice of the reference point $\vec{\epsilon}^*$ for the Taylor expansion of the perturbative free energy. The partition function should not depend on this choice, and it can indeed be checked that the $\vec{\epsilon}^*$-derivative of its expression (4.106) vanishes to each order in N. This property is called **background independence**, and has analogs in general relativity, string theory, etc.

This does not mean that the reference filling fractions $\vec{\epsilon}^*$ can be chosen arbitrarily. For the Theta function $\Theta\left(2\pi i\vec{\mu} - NF_0^{(\vec{1})}(\vec{\epsilon}^*), \frac{1}{2\pi i}F_0^{(\vec{2})}(\vec{\epsilon}^*)\right)$ that appears in eq. (4.106) to remain finite as $N \to \infty$, we need $\Re F_0^{(\vec{1})}(\vec{\epsilon}^*) = O(\frac{1}{N})$, so that $\vec{\epsilon}^*$ should be close (at a distance $O(\frac{1}{N})$) to a minimum of $\Re F_0$.

Now, at a minimum of $\Re F_0$, we have

$$\forall\, i, \qquad \frac{\partial \Re F_0}{\partial \epsilon_i} = \Re \oint_{\mathcal{B}_i} y dx = 0 \,. \tag{4.118}$$

For any choice of real filling fractions, we moreover have

$$\forall\, i, \qquad \Re \oint_{\mathcal{A}_i} y dx \; = \Re 2\pi i \epsilon_i = 0 \,. \tag{4.119}$$

Combining these two equations, we obtain the vanishing of the real part of the integral of ydx on any contour,

$$\boxed{\forall \mathcal{C}, \qquad \Re \oint_{\mathcal{C}} y dx \; = 0} \,. \tag{4.120}$$

The spectral curve is then said to have the **Boutroux property**. Therefore, our expression for the partition function makes sense provided the reference spectral curve that corresponds to the reference point $\vec{\epsilon}^*$ is close to a Boutroux algebraic curve.

Chapter 5

Orthogonal polynomials and integrable systems

In this Chapter we use orthogonal polynomials for exactly computing matrix integrals. This also provides a link between matrix integrals and integrable systems.

5.1 Determinantal formulas

We consider a normal matrix integral (1.34), written as an integral over eigenvalues

$$\mathcal{Z} = \frac{1}{N!} \int_{\gamma^N} d\lambda_0 \cdots d\lambda_{N-1} \ \Delta(\lambda_0, \cdots, \lambda_{N-1})^2 \prod_{i=0}^{N-1} e^{-V(\lambda_i)} \ , \tag{5.1}$$

where $\Delta(\lambda_0, \cdots \lambda_{N-1})$ is the Vandermonde determinant. The integration path γ can be arbitrary. Our analysis can be generalized to the cases $\beta = 1, 4$ of the Orthogonal and Symplectic ensembles. For later convenience we have introduced the overall normalization factor $\frac{1}{N!}$, and labelled the eigenvalues from 0 to $N-1$.

5.1.1 Matrix integrals

Our first aim is to separate the variables in the matrix integral $\mathcal{Z}$. To do this, we rewrite the Vandermonde determinant as a linear combination of monomials,

$$\Delta(\lambda_0, \cdots, \lambda_{N-1}) = \det_{i,j=0,\cdots N-1} \lambda_i^j = \sum_{\sigma \in \mathfrak{S}_N} (-1)^\sigma \prod_{i=0}^{N-1} \lambda_i^{\sigma(i)} \ , \tag{5.2}$$

where $\mathfrak{S}_N$ is the group of permutations of N elements, called the symmetric group, and $(-1)^\sigma$ is the signature of $\sigma \in \mathfrak{S}_N$. This leads to

$$\mathcal{Z} = \frac{1}{N!} \sum_{\sigma, \sigma' \in \mathfrak{S}_N} (-1)^{\sigma\sigma'} \prod_{i=0}^{N-1} \int_\gamma d\lambda_i \ \lambda_i^{\sigma(i)} \lambda_i^{\sigma'(i)} e^{-V(\lambda_i)} \ . \tag{5.3}$$

We have indeed separated the variables λ_i. We notice that the resulting one-dimensional integrals are examples of the scalar product for the measure $e^{-V(\lambda)} d\lambda$,

$$\langle f | g \rangle = \int_\gamma d\lambda \ e^{-V(\lambda)} f(\lambda) g(\lambda) \ , \tag{5.4}$$

where in our case the functions f and g are monomials of the type

$$m_i(\lambda) = \lambda^i \ . \tag{5.5}$$

We thus compute

$$\mathcal{Z} = \frac{1}{N!} \sum_{\sigma,\sigma' \in \mathfrak{S}_N} (-1)^{\sigma\sigma'} \prod_{i=0}^{N-1} \big\langle m_{\sigma(i)} \big| m_{\sigma'(i)} \big\rangle = \sum_{\sigma \in \mathfrak{S}_N} (-1)^{\sigma} \prod_{i=0}^{N-1} \big\langle m_{\sigma(i)} \big| m_i \big\rangle \ , \tag{5.6}$$

and we end up with a **determinantal formula** for our matrix integral,

$$\mathcal{Z} = \det_{i,j=0,\cdots N-1} \langle m_i | m_j \rangle \ . \tag{5.7}$$

This is a **Hankel determinant**, that is a determinant of the type

$$\mathcal{Z} = \det_{i,j=0,\cdots N-1} M_{i+j} \qquad , \quad M_k = \int_\gamma d\lambda \ \lambda^k e^{-V(\lambda)} \ . \tag{5.8}$$

By permuting the columns $j \to N-1-j$, this can be rewritten in terms of a **Toeplitz determinant** whose entries are constant on diagonals, i.e. depend only on $i-j$:

$$\mathcal{Z} = (-1)^{\frac{N(N-1)}{2}} \det_{i,j=0,\cdots N-1} M_{N-1+i-j} \ . \tag{5.9}$$

In our determinantal formula, the monomials m_i can be replaced with arbitrary monic polynomials p_i of the same degrees. Actually, introducing two families of such polynomials $p_i(\lambda) = \lambda^i + \cdots$ and $\tilde{p}_i(\lambda) = \lambda^i + \cdots$, and their scalar products

$$\boxed{H_{i,j} = \langle \tilde{p}_i | p_j \rangle} \ , \tag{5.10}$$

we have

$$\boxed{\mathcal{Z} = \det_{i,j=0,\cdots N-1} H_{i,j}} \ . \tag{5.11}$$

5.1.2 Joint eigenvalue distributions

We will now study the **joint eigenvalue distributions**

$$\boxed{R_k(\lambda_0, \cdots \lambda_{k-1}) = \frac{1}{\mathcal{Z}} \frac{1}{N!} \int_{\gamma^{N-k}} d\lambda_k \cdots d\lambda_{N-1} \ \Delta(\lambda_0, \cdots \lambda_{N-1})^2 \prod_{i=0}^{N-1} e^{-V(\lambda_i)}} \ . \tag{5.12}$$

In particular,

- $R_0 = 1$,
- $R_1(\lambda) = \rho(\lambda)$ is the density of eigenvalues,
- R_2 is related to the connected two-point function ρ_2 by

$$\rho_2(\lambda_1, \lambda_2) = N(N-1) R_2(\lambda_1, \lambda_2) + N\delta(\lambda_1 - \lambda_2) R_1(\lambda_1) - N^2 R_1(\lambda_1) R_1(\lambda_2) \ . \tag{5.13}$$

- R_N is (up to a normalization) the integrand of our matrix integral $\mathcal{Z}$.

We first remark that R_N can be expressed in terms of three determinant factors: one from the determinantal formula (5.11) for $\mathcal{Z}$, and two from the Vandermonde determinants,

$$R_N(\lambda_0, \cdots \lambda_{N-1}) = \frac{1}{N!} \det H^{-1} \det_{i,j=0,\cdots N-1} p_j(\lambda_i) \det_{i',j'=0,\cdots N-1} \tilde{p}_{j'}(\lambda_{i'}) \prod_{i=0}^{N-1} e^{-V(\lambda_i)} \ . \tag{5.14}$$

The product of these three determinants can be rewritten as the determinant of a product matrix, and we find

$$R_N(\lambda_0, \cdots \lambda_{N-1}) = \frac{1}{N!} \det_{i,j=0,\cdots N-1} K(\lambda_i, \lambda_j) \ , \tag{5.15}$$

where we introduce

$$\boxed{K(\lambda, \tilde{\lambda}) = \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} p_k(\lambda) \left(H^{-1}\right)_{k,k'} \tilde{p}_{k'}(\tilde{\lambda}) e^{-\frac{1}{2}V(\lambda)} e^{-\frac{1}{2}V(\tilde{\lambda})}} \ . \tag{5.16}$$

The kernel K enjoys the following properties:

- K is independent from the choice of the families of polynomials p_k and $\tilde{p}_k$. Choosing another family of polynomials p_k indeed amounts to multiplying the vector $p = (p_0, \cdots p_{N-1})$ and the matrix H on the right with the same matrix, so that $\sum_{k=0}^{N-1} p_k \left(H^{-1}\right)_{k,k'}$ is unchanged.
- $\int_\gamma d\lambda \ K(\lambda, \lambda) = \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} \left(H^{-1}\right)_{k,k'} H_{k',k} = N$,
- K is a **self-reproducing kernel**, which obeys

$$K \star K = K \ , \tag{5.17}$$

where we define the star product of two functions $f(\lambda^1, \cdots \lambda^p)$ and $g(\mu^1, \cdots \mu^q)$ of one or more variables as

$$(f \star g)(\lambda^1, \cdots \lambda^{p-1}, \mu^2, \cdots \mu^q) = \int_\gamma d\lambda \ f(\lambda^1, \cdots \lambda^{p-1}, \lambda) g(\lambda, \mu^2, \cdots \mu^q) \ . \tag{5.18}$$

Let us now compute the next eigenvalue distribution

$$R_{N-1}(\lambda_0, \cdots \lambda_{N-2}) = \frac{1}{N!} \int_\gamma d\lambda_{N-1} \ R_N(\lambda_0, \cdots \lambda_{N-1}) \ , \tag{5.19}$$

$$= \frac{1}{N!} \int_\gamma d\lambda_{N-1} \ \sum_{\sigma \in \mathfrak{S}_N} (-1)^\sigma K(\lambda_i, \lambda_{\sigma(i)}) \ . \tag{5.20}$$

Let us rewrite the permutations $\sigma \in \mathfrak{S}_N$ in terms of permutations $\sigma' \in \mathfrak{S}_{N-1}$ of the first $N-1$ elements. If $\sigma(N-1) = N-1$, then we write $\sigma = \sigma' \cdot (N-1)$ as the composition of $\sigma' \in \mathfrak{S}_{N-1}$ with the identity acting on the last element $N-1$. And if $\sigma(N-1) = i \neq N-1$, then we write $\sigma = \sigma' \cdot (N-1, i)$ as the composition of $\sigma' \in \mathfrak{S}_{N-1}$ with a transposition.

We thus find

$$N!R_{N-1} = \sum_{\sigma'\in\mathfrak{S}_{N-1}} (-1)^{\sigma'} \prod_{i=0}^{N-2} K(\lambda_i,\lambda_{\sigma(i)}) \int_\gamma d\lambda_{N-1}\, K(\lambda_{N-1},\lambda_{N-1})$$
$$- \sum_{i=0}^{N-2} \sum_{\sigma'\in\mathfrak{S}_{N-1}} (-1)^{\sigma'} \prod_{\substack{j=0\\ j\neq i}}^{N-2} K(\lambda_j,\lambda_{\sigma'(j)}) \int_\gamma d\lambda_{N-1}\, K(\lambda_i,\lambda_{N-1}) K(\lambda_{N-1},\lambda_{\sigma'(i)})\,, \tag{5.21}$$

$$= N \det_{i,j=0,\cdots N-2} K(\lambda_i,\lambda_j) - (N-1) \det_{i,j=0,\cdots N-2} K(\lambda_i,\lambda_j)\,, \tag{5.22}$$

$$= \det_{i,j=0,\cdots N-2} K(\lambda_i,\lambda_j)\,. \tag{5.23}$$

So R_{N-1} is also given by a determinantal formula. Similarly, we can find R_{N-2} from R_{N-1} by integrating over λ_{N-2}. Iterating, we obtain **Dyson's theorem** [1] – determinantal formulas for all the joint eigenvalue distributions,

$$\boxed{R_k(\lambda_0,\cdots\lambda_{k-1}) = \frac{(N-k)!}{N!} \det_{i,j=0,\cdots k-1} K(\lambda_i,\lambda_j)}\,. \tag{5.24}$$

In particular, the eigenvalue density and the connected two-point function can be written as

$$\rho(\lambda) = \frac{1}{N} K(\lambda,\lambda)\ , \tag{5.25}$$

$$\rho_2(\lambda_1,\lambda_2) = -K(\lambda_2,\lambda_1)K(\lambda_1,\lambda_2) + \delta(\lambda_1-\lambda_2)K(\lambda_1,\lambda_1)\ . \tag{5.26}$$

More generally, any disconnected k-point correlation function is given by a determinant $\det_{i,j=0,\cdots k-1} K(\lambda_i,\lambda_j)$, plus delta-function terms.

5.1.3 Spacing distributions

Given a set $I\subset\gamma$, let $\mathcal{E}(I)$ be the probability that there is no eigenvalue in I. From this probability we can in particular deduce the **spacing distribution**, that is the probability that $[a,b]$ is an interval between two consecutive eigenvalues a and b,

$$\boxed{p_{\text{spacing}}(a,b) = -\frac{1}{N(N-1)} \frac{\partial^2}{\partial a \partial b} \mathcal{E}([a,b])}\,. \tag{5.27}$$

In the case $I=[a,\infty)$, we find the probability that a is the largest eigenvalue:

$$\boxed{p_{\text{largest}}(a) = -\frac{1}{N}\frac{\partial}{\partial a}\mathcal{E}([a,\infty))}\,. \tag{5.28}$$

The probability $\mathcal{E}(I)$ is given by the matrix integral

$$\mathcal{E}(I) = \frac{1}{\mathcal{Z}}\frac{1}{N!} \int_{(\gamma-I)^N} d\lambda_0\cdots d\lambda_{N-1}\ \Delta(\lambda_0,\cdots,\lambda_{N-1})^2 \prod_{i=0}^{N-1} e^{-V(\lambda_i)}\ . \tag{5.29}$$

This is very similar to the matrix integral $\mathcal{Z}$ itself, with the integration contour γ replaced with $\gamma - I$. Explicitly writing the γ-dependences of the scalar product and of the matrix $(H_\gamma)_{i,j} = \det_{i,j=0,\cdots N-1} \langle \tilde{p}_i | p_j\rangle_\gamma$ (5.10), we find

$$\mathcal{E}(I) = \frac{\det H_{\gamma-I}}{\det H_\gamma} = \det(\mathrm{Id} - A) \qquad \text{with} \qquad A = H_\gamma^{-1} H_I\ . \tag{5.30}$$

It turns out that the transposed matrix A^T is the matrix (in a particular basis) of the operator K_I that acts on functions on I as

$$\boxed{K_I(f)(\lambda) = (K \star_I f)(\lambda) = \int_I d\tilde{\lambda}\ K(\lambda, \tilde{\lambda}) f(\tilde{\lambda})} \,. \tag{5.31}$$

The particular basis is made of the functions $p_j e^{-\frac{1}{2}V}$, and we indeed have

$$K_I(p_j e^{-\frac{1}{2}V})(\lambda) = \int_I d\tilde{\lambda} \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} p_k(\lambda) \left(H_\gamma^{-1}\right)_{k,k'} \tilde{p}_{k'}(\tilde{\lambda}) p_j(\tilde{\lambda}) e^{-\frac{1}{2}V(\lambda)} e^{-V(\tilde{\lambda})} \ , \tag{5.32}$$

$$= \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} p_k(\lambda) e^{-\frac{1}{2}V(\lambda)} \left(H_\gamma^{-1}\right)_{k,k'} (H_I)_{k',j} \ , \tag{5.33}$$

$$= \left(\sum_{k=0}^{N-1} p_k e^{-\frac{1}{2}V} A_{k,j} \right)(\lambda) \ . \tag{5.34}$$

This implies that $\mathcal{E}(I)$ can be written as the Fredholm determinant

$$\boxed{\mathcal{E}(I) = \det\left(\mathrm{Id} - K_I\right)} \,. \tag{5.35}$$

For example, in the large N bulk microscopic limit, the spacing distribution is the (second derivative of the) Fredholm determinant of the sine kernel. In the edge microscopic limit, the Tracy–Widom law for the largest eigenvalue is the (derivative of the) Fredholm determinant of the Airy kernel.

5.2 Orthogonal polynomials

Our formulas for matrix integrals and eigenvalue distributions, which involve scalar products of polynomials, would simplify if the polynomials in question were orthogonal to one another. Let us assume that $p_k = \tilde{p}_k$ for a family of **orthogonal polynomials**, so that

$$\boxed{H_{k,k'} = \langle p_k | p_{k'} \rangle = h_k \delta_{k,k'}} \,, \tag{5.36}$$

for some numbers $h_k \in \mathbb{C}$. (We cannot set $h_k = 1$ because our polynomials are assumed to be monic, and cannot be arbitrarily renormalized.) The partition function is then

$$\boxed{\mathcal{Z}_N = \prod_{k=0}^{N-1} h_k} \,, \tag{5.37}$$

where we now indicate the matrix size N in $\mathcal{Z}_N$. Moreover, the self-reproducing kernel can then be written as

$$\boxed{K_N(\lambda, \tilde{\lambda}) = \sum_{k=0}^{N-1} \psi_k(\lambda) \psi_k(\tilde{\lambda})} \,, \tag{5.38}$$

where we define

$$\boxed{\psi_k(\lambda) = \frac{p_k(\lambda)}{\sqrt{h_k}} e^{-\frac{1}{2}V(\lambda)} \qquad \text{so that} \qquad \int_\gamma d\lambda\ \psi_k(\lambda) \psi_{k'}(\lambda) = \delta_{k,k'}} \,. \tag{5.39}$$

The existence and uniqueness of an orthogonal family of monic polynomials is equivalent to the non-degeneracy of the scalar product, as follows from the Gram–Schmidt construction. Due to the determinantal formula, this is also equivalent to $\forall N\ \mathcal{Z}_N \neq 0$. This condition is fulfilled if $\gamma = \mathbb{R}$ and the potential is real, in which case $\forall N\ \mathcal{Z}_N > 0$, and the existence of the orthogonal family implies that the density of eigenvalues is strictly positive,

$$N\rho(\lambda) = K_N(\lambda,\lambda) = \sum_{k=0}^{N-1} \psi_k(\lambda)^2 \geq \psi_0(\lambda)^2 = \frac{e^{-V(\lambda)}}{h_0} > 0\ . \tag{5.40}$$

Given a potential, the scalar product depends polynomially on the contour $\gamma = \sum_{i=1}^{d} c_i\gamma_i$, and is therefore degenerate on an enumerable union of algebraic submanifolds, which is a subset of measure zero of the space $\mathbb{C}^d$ of contours. Orthogonal polynomials therefore exist in a dense subset of this space.

5.2.1 How to compute orthogonal polynomials

In the particularly simple case of a Gaussian potential $V(\lambda) = \frac{1}{2}\lambda^2$, an orthogonal family is provided by the **Hermite polynomials**,

$$p_k(\lambda) = (-1)^k e^{\frac{\lambda^2}{2}} \left(\frac{\partial}{\partial\lambda}\right)^k e^{-\frac{\lambda^2}{2}}\ . \tag{5.41}$$

Let us see the expressions and properties of the orthogonal polynomials in more general cases.

Heine's formula

Expressions for the orthogonal polynomials are given by **Heine's formula**,

$$\boxed{p_k(\lambda) = \Big\langle \det(\lambda - M) \Big\rangle_{k\text{-matrix}}}\ , \tag{5.42}$$

or more explicitly

$$p_k(\lambda) = \frac{\int_{\gamma^k} d\lambda_0 \cdots d\lambda_{k-1}\ \Delta(\lambda_0,\cdots\lambda_{k-1})^2 \prod_{i=0}^{k-1}(\lambda-\lambda_i)e^{-V(\lambda_i)}}{\int_{\gamma^k} d\lambda_0 \cdots d\lambda_{k-1}\ \Delta(\lambda_0,\cdots\lambda_{k-1})^2 \prod_{i=0}^{k-1} e^{-V(\lambda_i)}}\ . \tag{5.43}$$

The polynomials p_k that are defined by Heine's formula are manifestly monic. To show that p_k is orthogonal to p_j for $j < k$, we compute

$$\langle p_k | p_j \rangle \propto \int_{\gamma^{k+1}} d\lambda d\lambda_0 \cdots d\lambda_{k-1}\ \Delta(\lambda_0,\cdots\lambda_{k-1})^2 p_j(\lambda) \prod_{i=0}^{k-1}(\lambda-\lambda_i)e^{-V(\lambda_i)}\ . \tag{5.44}$$

The presence of the factor $\Delta(\lambda_0, \cdots \lambda_{k-1}) \prod_{i=0}^{k-1}(\lambda - \lambda_i) = \Delta(\lambda_0, \cdots \lambda_{k-1}, \lambda)$ allows us to antisymmetrize the rest of the integrand:

$$\Delta(\lambda_0, \cdots \lambda_{k-1}) p_j(\lambda) = \det \begin{pmatrix} 1 & 1 & \cdots & 1 & 0 \\ \lambda_0 & \lambda_1 & \cdots & \lambda_{k-1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_0^{k-1} & \lambda_1^{k-1} & \cdots & \lambda_{k-1}^{k-1} & 0 \\ 0 & 0 & \cdots & 0 & p_j(\lambda) \end{pmatrix} \tag{5.45}$$

$$= \det \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ \lambda_0 & \lambda_1 & \cdots & \lambda_{k-1} & \lambda \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_0^{k-1} & \lambda_1^{k-1} & \cdots & \lambda_{k-1}^{k-1} & \lambda^{k-1} \\ 0 & 0 & \cdots & 0 & p_j(\lambda) \end{pmatrix} \tag{5.46}$$

$$\underset{\text{antisymmetrize}}{\to} \det \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ \lambda_0 & \lambda_1 & \cdots & \lambda_{k-1} & \lambda \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_0^{k-1} & \lambda_1^{k-1} & \cdots & \lambda_{k-1}^{k-1} & \lambda^{k-1} \\ p_j(\lambda_0) & p_j(\lambda_1) & \cdots & p_j(\lambda_{k-1}) & p_j(\lambda) \end{pmatrix} \tag{5.47}$$

$$= 0 \quad \text{for} \quad j < k \ . \tag{5.48}$$

So Heine's formula does yield orthogonal polynomials. But this formula is still itself a matrix integral. We will now see more explicit expressions of the same orthogonal polynomials.

Orthogonal polynomials as Hankel determinants

Another formula for orthogonal polynomials is

$$p_k(\lambda) = \frac{\det \begin{pmatrix} M_0 & M_1 & M_2 & \dots & M_{k-1} & 1 \\ M_1 & M_2 & & & M_k & \lambda \\ M_2 & & & & M_{k+1} & \lambda^2 \\ \vdots & & & & \vdots & \vdots \\ M_k & & & & M_{2k-1} & \lambda^k \end{pmatrix}}{\det \begin{pmatrix} M_0 & M_1 & M_2 & \dots & M_{k-1} \\ M_1 & M_2 & & & M_k \\ M_2 & & & & M_{k+1} \\ \vdots & & & & \vdots \\ M_{k-1} & & & & M_{2k-2} \end{pmatrix}} \ , \tag{5.49}$$

where $M_j = \int_\gamma \lambda^j e^{-V(\lambda)}$ are the moments. Indeed this expression is a monic polynomial of degree k, and satisfies

$$\int_\gamma d\lambda\, \lambda^j p_k(\lambda) e^{-V(\lambda)} = \frac{\det \begin{pmatrix} M_0 & M_1 & M_2 & \dots & M_{k-1} & M_j \\ M_1 & M_2 & & & M_k & M_{j+1} \\ M_2 & & & & M_{k+1} & M_{j+2} \\ \vdots & & & & \vdots & \vdots \\ M_k & & & & M_{2k-1} & M_{j+k} \end{pmatrix}}{\det \begin{pmatrix} M_0 & M_1 & M_2 & \dots & M_{k-1} \\ M_1 & M_2 & & & M_k \\ M_2 & & & & M_{k+1} \\ \vdots & & & & \vdots \\ M_{k-1} & & & & M_{2k-2} \end{pmatrix}} , \tag{5.50}$$

which vanishes if $j \leq k-1$. This proves the formula (5.49) for the orthogonal polynomials. This formula is more explicit than Heine's formula, but still involves complicated determinants. We will shortly do better.

Recursion relation

We will use the recursion relation

$$m_1 p_k = \sum_{j=0}^{k+1} \hat{Q}_{k,j} p_j \quad \text{with} \quad \hat{Q}_{k,k+1} = 1 \ , \tag{5.51}$$

where $m_1(x) = x$ is still the monomial of degree one. This relation exists and is uniquely defined, because $m_1 p_k$ is a monic polynomial of degree $k+1$. Inserting this recursion relation in the identity $\langle m_1 p_k | p_j \rangle = \langle p_k | m_1 p_j \rangle$, we obtain the following identity for the unknown coefficients $\hat{Q}_{k,j}$,

$$h_j \hat{Q}_{k,j} = h_k \hat{Q}_{j,k} \ . \tag{5.52}$$

In particular, if $j < k-1$ then $\hat{Q}_{k,j} = 0$, so the right-hand side of the recursion relation has only three non-vanishing terms. Using the functions ψ_k (5.39) instead of the polynomials p_k, the three-term recursion relation becomes

$$\boxed{m_1 \psi_k = Q_{k,k+1} \psi_{k+1} + Q_{k,k} \psi_k + Q_{k,k-1} \psi_{k-1}} \ , \tag{5.53}$$

where we define

$$Q_{j,k} = Q_{k,j} = \sqrt{\frac{h_j}{h_k}} \hat{Q}_{k,j} \ , \tag{5.54}$$

The infinite-dimensional, symmetric, tridiagonal matrix Q is called the **Jacobi matrix**, and can be written as

$$Q = \begin{pmatrix} S_0 & \gamma_1 & 0 & 0 & \cdots \\ \gamma_1 & S_1 & \gamma_2 & 0 & \ddots \\ 0 & \gamma_2 & S_2 & \gamma_3 & \ddots \\ 0 & 0 & \gamma_3 & S_3 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \quad \text{with} \quad \boxed{\gamma_k = Q_{k,k-1} = \sqrt{\frac{h_k}{h_{k-1}}}} \quad \text{and} \quad \boxed{S_k = Q_{k,k}} \ . \tag{5.55}$$

The Jacobi matrix Q is useful for computing expectation values of single-trace operators $\operatorname{Tr} f(M)$. For any polynomial function f, we indeed have

$$\boxed{\left\langle \operatorname{Tr} f(M) \right\rangle_{N\text{-matrix}} = \operatorname{Tr} \Pi_N f(Q)} \,, \tag{5.56}$$

where $\Pi_N = \begin{pmatrix} \mathrm{Id}_{N\times N} & 0_{\infty\times N} \\ 0_{N\times\infty} & 0_{\infty\times\infty} \end{pmatrix}$ is the projector on the first N components. To prove this, we will directly compute the relevant integral over eigenvalues. Writing the Vandermonde determinant as $\Delta(\lambda_0,\dots,\lambda_{N-1}) = \sum_{\sigma\in\mathfrak{S}_N}(-1)^\sigma \prod_{i=0}^{N-1} p_{\sigma(i)}(\lambda_i)$, we find

$$\left\langle \operatorname{Tr} f(M) \right\rangle_{N\text{-matrix}} = \frac{1}{N!\prod_{k=0}^{N-1} h_k} \sum_{i=0}^{N-1} \sum_{\sigma,\sigma'\in\mathfrak{S}_N} \left\langle f p_{\sigma(i)} \middle| p_{\sigma'(i)} \right\rangle \prod_{j\neq i} \left\langle p_{\sigma(j)} \middle| p_{\sigma'(j)} \right\rangle \,, \tag{5.57}$$

$$= \frac{1}{N!} \sum_{\sigma\in\mathfrak{S}_N} \sum_{i=0}^{N-1} \frac{1}{h_{\sigma(i)}} \left\langle f p_{\sigma(i)} \middle| p_{\sigma(i)} \right\rangle = \sum_{k=0}^{N-1} \frac{1}{h_k} \left\langle f p_k | p_k \right\rangle \,. \tag{5.58}$$

Replacing the polynomials p_k with the functions ψ_k, and using the recursion relation $f\psi_k = \sum_j f(Q)_{k,j}\psi_j$, we obtain eq. (5.56).

The orthogonal polynomials p_k can be expressed in terms of submatrices of Q by the **determinantal formula**

$$\boxed{p_k(\lambda) = \det_{k\times k \text{ submatrix}} (\lambda - Q)} \,. \tag{5.59}$$

The polynomials on both sides of this equation indeed have the same zeros: if $p_k(\lambda) = 0$, then the vector $(p_0(\lambda), \cdots p_{k-1}(\lambda))$ is killed by the matrix $(\lambda - Q)_{k\times k \text{ submatrix}}$, due to the recursion relation – and this implies $\det_{k\times k \text{ submatrix}} (\lambda - Q) = 0$. Combining this equation with Heine's formula, we obtain $\langle \det(\lambda - M) \rangle_{k\text{-matrix}} = \det_{k\times k \text{ submatrix}} (\lambda - Q)$.

The matrix Q can also be used for writing the self-reproducing kernel,

$$\boxed{K_{N+1}(\lambda,\mu) = \frac{e^{-\frac{1}{2}V(\lambda)} e^{-\frac{1}{2}V(\mu)}}{h_N} \det_{N\times N \text{ submatrix}} (\lambda - Q)(\mu - Q)} \,. \tag{5.60}$$

Let us prove this by induction on N. Assuming this holds for K_N, and denoting $(\lambda - Q)_N$ the upper left $N \times N$ block of $(\lambda - Q)$, we have

$$\Big((\lambda - Q)(\mu - Q) \Big)_N = (\lambda - Q)_N (\mu - Q)_N + \gamma_N^2 \operatorname{diag}_{N\times N}(0,0,\dots,0,1) \,. \tag{5.61}$$

It follows that

$$\det_{N\times N} (\lambda - Q)(\mu - Q) = \det (\lambda - Q)_N (\mu - Q)_N + \gamma_N^2 \det_{N-1\times N-1} (\lambda - Q)_N (\mu - Q)_N \,, \tag{5.62}$$

$$= p_N(\lambda) p_N(\mu) + \gamma_N^2 h_{N-1} e^{\frac{1}{2}V(\lambda)} e^{\frac{1}{2}V(\mu)} K_N(\lambda,\mu) \,, \tag{5.63}$$

$$= h_N e^{\frac{1}{2}V(\lambda)} e^{\frac{1}{2}V(\mu)} K_{N+1}(\lambda,\mu) \,, \tag{5.64}$$

where we used the expression (5.38) of the kernel in terms of the orthogonal polynomials.

Derivatives of orthogonal polynomials

In the same way as we wrote a recursion relation for the multiplication of our polynomials with the monomial m_1, we can write a recursion relation for taking derivatives of our polynomials p_k – or equivalently of the orthonormal functions ψ_k. The second option turns out to be more convenient for normalizations, although it is also more complicated, since ψ_k is not a polynomial. Writing

$$d = \deg V' \ , \tag{5.65}$$

$V'p_k$ is a polynomial of degree $k+d$, and we must have

$$\boxed{\psi_k' = \sum_{j=0}^{k+d} P_{k,j}\psi_j} \ , \tag{5.66}$$

for some coefficients $P_{k,j}$. Due to the identity $\int_\gamma \psi_k'\psi_j = -\int_\gamma \psi_k\psi_j'$, the matrix P must be antisymmetric, and its nonvanishing coefficients reside on $2d$ bands:

$$P = \begin{pmatrix} 0 & P_{0,1} & \cdots & P_{0,d} & 0 & 0 & \cdots \\ -P_{0,1} & 0 & P_{1,2} & \cdots & P_{1,d+1} & 0 & \cdots \\ \vdots & -P_{1,2} & 0 & P_{2,3} & \cdots & P_{2,d+2} & \ddots \\ -P_{0,d} & \vdots & -P_{2,3} & 0 & P_{3,4} & \ddots & \ddots \\ 0 & -P_{1,d+1} & \vdots & -P_{3,4} & 0 & P_{4,5} & \ddots \\ 0 & 0 & -P_{2,d+2} & \ddots & -P_{4,5} & 0 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \tag{5.67}$$

Now, using definition of ψ_k, we have

$$\frac{1}{\sqrt{h_k}} p_k' e^{-\frac{1}{2}V} = \psi_k' + \frac{1}{2}V'\psi_k \ , \tag{5.68}$$

$$= \sum_{j=0}^{k+d} \left(P + \frac{1}{2}V'(Q) \right)_{kj} \psi_j \ . \tag{5.69}$$

However, since $\deg p_k' = k-1$, the left-hand side of this equation must be a combination of $\psi_0, \cdots \psi_{k-1}$. This shows that the infinite matrix $P + \frac{1}{2}V'(Q)$ must be strictly lower triangular. Denoting M_+, M_0 and M_- the upper triangular, diagonal and lower triangular parts of a possibly infinite matrix M, this implies

$$P_+ = -\frac{1}{2}V'(Q)_+ \ . \tag{5.70}$$

By antisymmetry of P and symmetry of Q, this is equivalent to

$$P_- = \frac{1}{2}V'(Q)_- \ , \tag{5.71}$$

and thus

$$\boxed{P = -\frac{1}{2}\Big(V'(Q)_+ - V'(Q)_- \Big)} \ . \tag{5.72}$$

This determines P once we know Q. Moreover, from $p'_k = kp_{k-1} + \dots$ we can determine the diagonal and subdiagonal coefficients of $P + \frac{1}{2}V'(Q)$,

$$\forall\, k\,, \qquad \begin{cases} \left(P + \frac{1}{2}V'(Q)\right)_{k,k-1} = \frac{k}{\gamma_k}\ , \\ \left(P + \frac{1}{2}V'(Q)\right)_{k,k} = 0\ . \end{cases} \tag{5.73}$$

Using eq. (5.72), this can be rewritten in terms of Q alone as

$$\boxed{\forall\, k\,, \qquad \begin{cases} \Big(V'(Q)\Big)_{k,k-1} = \frac{k}{\gamma_k}\ , \\ \Big(V'(Q)\Big)_{k,k} = 0\ . \end{cases}} \tag{5.74}$$

These two equations lead to recursion relations for $S_k = Q_{k,k}$ and $\gamma_k = Q_{k,k-1}$, which determine them. Let us give their values in two examples:

- Gaussian potential $V'(Q) = Q$: we find $S_k = S_0$ and $\gamma_k = \sqrt{k}$.
- Quartic potential $V'(Q) = \frac{1}{t}(Q - Q^3)$: we find $S_k = 0$, and eq. (5.74) implies a recursion relation for $R_k = \gamma_k^2$,

$$\frac{R_k}{t}(1 - R_{k-1} - R_k - R_{k+1}) = k\ . \tag{5.75}$$

Finally, let us mention an alternative derivation of the equations (5.74). Instead of using the relation $p'_k = kp_{k-1} + \cdots$, we could use the **string equation**

$$\boxed{[Q, P] = \mathrm{Id}}\ , \tag{5.76}$$

which follows from the definitions of P and Q as the matrices of the derivative and multiplication with m_1 respectively. Here we will check that the results (5.72) - (5.74) for P and Q imply the string equation. Since $P + \frac{1}{2}V'(Q)$ is strictly lower triangular and Q has at most one band above diagonal, $[Q, P] = [Q, P + \frac{1}{2}V'(Q)]$ must be lower triangular. The same reasoning with $P - \frac{1}{2}V'(Q)$ shows that $[Q, P]$ is also upper triangular, and therefore diagonal. And its diagonal coefficients are

$$\left[Q, P + \frac{1}{2}V'(Q)\right]_{k,k} = Q_{k,k+1}\left(P + \frac{1}{2}V'(Q)\right)_{k+1,k} - \left(P + \frac{1}{2}V'(Q)\right)_{k,k-1} Q_{k-1,k}\ , \tag{5.77}$$

$$= \gamma_{k+1}\frac{k+1}{\gamma_{k+1}} - \frac{k}{\gamma_k}\gamma_k = 1\ . \tag{5.78}$$

5.2.2 Hilbert transforms of orthogonal polynomials

We will shortly see that ψ_k obeys a second-order differential equation. In order to construct another linearly independent solution, it is useful to introduce the **Hilbert transform** of ψ_k,

$$\boxed{\varphi_k(x) = e^{\frac{1}{2}V(x)} \int_\gamma dx' \frac{\psi_k(x')}{x - x'} e^{-\frac{1}{2}V(x')}}\ . \tag{5.79}$$

Let us study the properties of φ_k, starting with its behaviour at infinity. Expanding $\frac{1}{x-x'} = \sum_{j=0}^{\infty} \frac{x'^j}{x^{j+1}}$, and noticing that $\int_\gamma dx' \psi_k(x') x'^j e^{-\frac{1}{2}V(x')} = 0$ if $j < k$, we find

$$\boxed{\varphi_k(x) \underset{x\to\infty}{=} \frac{\sqrt{h_k}}{x^{k+1}} e^{\frac{1}{2}V(x)} \left(1 + O\left(\frac{1}{x}\right)\right)} \,. \tag{5.80}$$

We accept that the functions φ_k obey the Heine-type formulas

$$\varphi_{k-1}(x) = \sqrt{h_{k-1}} e^{\frac{1}{2}V(x)} \left\langle \frac{1}{\det(x-M)} \right\rangle_{k-\text{matrix}} \,. \tag{5.81}$$

The functions φ_k obey recursion and derivation relations, which only differ from the corresponding relations for ψ_k for the first few values of k:

$$x\varphi_k(x) = \sum_{j=k-1}^{k+1} Q_{k,j}\varphi_j(x) + \frac{\delta_{k,0}}{\psi_0(x)} \,, \tag{5.82}$$

$$= \sum_{j=k-1}^{k+1} Q_{k,j}\varphi_j(x) \qquad \text{if} \quad k \geq 1 \,, \tag{5.83}$$

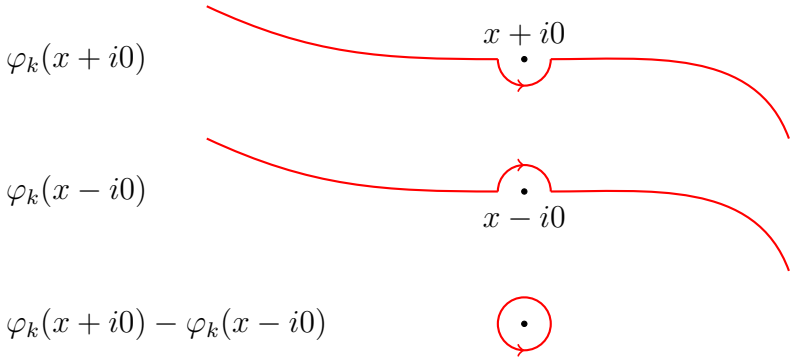
$$\varphi_k'(x) = \sum_{j=k-d}^{k+d} P_{k,j}\varphi_j(x) + \frac{e^{\frac{1}{2}V(x)}}{2} \int_\gamma dx' \psi_k(x') \frac{V'(x) - V'(x')}{x-x'} e^{-\frac{1}{2}V(x')} \,, \tag{5.84}$$

$$= \sum_{j=k-d}^{k+d} P_{k,j}\varphi_j(x) \qquad \text{if} \quad k \geq d \,. \tag{5.85}$$

Notice that φ_k is discontinuous on the integration domain γ. More specifically, if $\gamma = \sum_i c_i \gamma_i$ is a linear combination of paths with complex coefficients c_i, then the behaviour of φ_k near $x \in \gamma_i$ is

$$\boxed{x \in \gamma_i \ : \qquad \varphi_k(x+i0) - \varphi_k(x-i0) = -2\pi i c_i \psi_k(x)} \,, \tag{5.86}$$

as can be deduced from the small deformations of γ_i which are needed for accommodating the points $x \pm i0$:



Finally, let us show that φ_k is given by the determinantal formula

$$\boxed{\varphi_{k-1}(x) = \sqrt{h_{k-1}} e^{\frac{1}{2}V(x)} \det_{k\times k \text{ submatrix}} \frac{1}{x-Q}} \,. \tag{5.87}$$

(Compare with the determinantal formula for the orthogonal polynomial p_k.) To prove this, we need to study matrix inverses of $x - Q$, starting with

$$\left(\frac{1}{x-Q}\right)_{j,k} = \int_\gamma dx' \frac{\psi_j(x')\psi_k(x')}{x-x'} \ , \tag{5.88}$$

which is a bilateral inverse of $x - Q$ as follows from the recursion relation (5.53). This can be rewritten as

$$\left(\frac{1}{x-Q}\right)_{j,k} = \psi_j(x)\varphi_k(x) + R_{j,k}(x) \ , \tag{5.89}$$

where we introduce

$$R_{j,k}(x) = -\frac{1}{\sqrt{h_j h_k}} \int_\gamma dx' \frac{p_j(x) - p_j(x')}{x-x'} p_k(x') e^{-V(x')} \ . \tag{5.90}$$

The matrix $R(x)$ is still a right inverse of $x - Q$, but is simpler than $\frac{1}{x-Q}$, as it is strictly lower triangular, and polynomial in x. More specifically, $R_{j,k}(x)$ is a polynomial of degree $j - k - 1$, with

$$R_{j,k}(x) \underset{x\to\infty}{=} -\sqrt{\frac{h_k}{h_j}} x^{j-k-1} \left(1 + O\left(\frac{1}{x}\right)\right) \ . \tag{5.91}$$

Actually, solving $(x - Q)R(x) = \mathrm{Id}$ for $R(x)$ lower triangular yields

$$\boxed{R_{n,m}(x) = \begin{cases} -\frac{1}{\gamma_{m+1}\dots\gamma_n} \det_{m+1\le i,j\le n-1}(x-Q)_{i,j} & \text{if } m+1 < n \ , \\ -\frac{1}{\gamma_n} & \text{if } m+1 = n \ , \\ 0 & \text{if } m \ge n \ . \end{cases}} \tag{5.92}$$

Similarly, the transpose $L(x) = R(x)^T$ is an upper triangular left inverse of $x - Q$. Introducing the infinite vectors $\vec{\psi} = (\psi_0, \psi_1, \cdots)^T$ and $\vec{\varphi} = (\varphi_0, \varphi_1, \cdots)^T$, we can write

$$\boxed{\frac{1}{x-Q} = \vec{\psi}(x)^T \vec{\varphi}(x) + R(x) = \vec{\varphi}(x)^T \vec{\psi}(x) + L(x)} \ . \tag{5.93}$$

The existence of these three different inverses illustrates the exotic properties of infinite matrices, which in particular do not form an associative algebra. For example,

$$0 = \frac{1}{x-Q}\left((x-Q)\vec{\psi}\right) \neq \left(\frac{1}{x-Q}(x-Q)\right)\vec{\psi} = \vec{\psi} \ . \tag{5.94}$$

The determinantal formula (5.87) for φ_{k-1} can now be proved by starting with the trivial relation

$$\varphi_{k-1}(x) = \sqrt{h_{k-1}} e^{\frac{1}{2}V(x)} (-1)^{k-1} \psi_0(x) \varphi_{k-1}(x) \prod_{j=1}^{k-1} R_{j,j-1} \ , \tag{5.95}$$

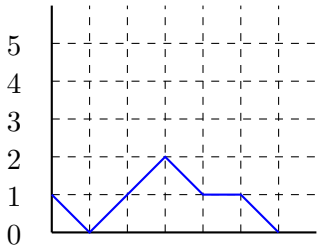
and rewriting the right-hand side as the desired minor of $\frac{1}{x-Q}$, with the help of eq. (5.93).

5.2.3 Motzkin paths and continuous fractions

We have found that the orthogonal polynomials (5.59) and the moments of our matrix model (5.56) can be expressed in terms of an infinite matrix Q, where however $|i - j| > 1 \Rightarrow Q_{i,j} = 0$. So a sum of the type $\sum_i Q_{i,j} \cdots$ has only three non-vanishing terms, which suggests that it can have a combinatorial interpretation in terms of paths with three possible directions, called Motzkin paths.

Motzkin paths

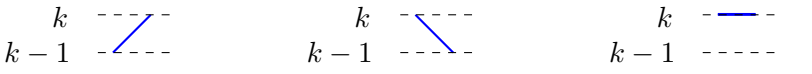
A **Motzkin path** is a path in the lattice $\mathbb{Z}_+ \times \mathbb{Z}_+$, made of edges with three possible directions: . For example,



We assign a weight to each edge depending on its height and direction, and the weight of a path is a product of the weights of its edges,

$$w(\mathfrak{m}) = \prod_{e \in \{\text{edges}(\mathfrak{m})\}} w(e) \ . \tag{5.96}$$

In particular, the weights that correspond to a matrix Q such that $|i-j|>1 \Rightarrow Q_{i,j}=0$ are



$$w_Q(e) = Q_{k-1,k} \qquad w_Q(e) = Q_{k,k-1} \qquad w_Q(e) = Q_{k,k}$$

Combinatorial expressions

Let us write a moment in terms of the matrix Q, using eq. (5.56):

$$\Big\langle \operatorname{Tr} M^m \Big\rangle_{n\text{-matrix}} = \operatorname{Tr} Q^m \Pi_n = \sum_{i_m=0}^{n-1} \sum_{i_1,\dots,i_{m-1}=0}^{\infty} Q_{i_m,i_1} Q_{i_1,i_2} Q_{i_2,i_3} \dots Q_{i_{m-1},i_m} \ . \tag{5.97}$$

We now interpret each nonzero matrix element as the weight of an edge, and each product of m matrix elements as the weight of a Motzkin path of length m. If $\mathcal{M}_{k,l}^{(m)}$ is the set of Motzkin paths of length m from height k to height l, our moment has the combinatorial expression

$$\Big\langle \operatorname{Tr} M^m \Big\rangle_{n\text{-matrix}} = \sum_{k=0}^{n-1} \sum_{\mathfrak{m} \in \mathcal{M}_{k,k}^{(n)}} w_Q(\mathfrak{m}) \ . \tag{5.98}$$

The matrix elements of Q are given in eq. (5.55), and we find for example:

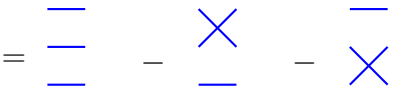
$$(Q^3)_{0,0} = \underset{S_0^3}{\text{⎽⎽⎽}} + \underset{\gamma_1^2 S_0}{\wedge} + \underset{S_0 \gamma_1^2}{\wedge} + \underset{\gamma_1 S_0 \gamma_1}{\cap}$$

The orthogonal polynomial $p_n(x)$ can be expressed as the determinant of a size n submatrix of $x-Q$. According to the Lindström–Gessel–Viennot lemma, this determinant is a weighted sum over a set of non-intersecting paths. (Here non-intersecting refers to the absence of intersection points on the lattice, whereas intersections outside the lattice

are allowed.) This amounts to summing over the set $\mathcal{M}^{(1)}_{\mathfrak{S}_n}$ of disconnected paths with n components of length 1 that correspond to permutations of $\{0,\dots,n-1\}$,

$$p_n(x) = \sum_{\mathfrak{m}\in\mathcal{M}^{(1)}_{\mathfrak{S}_n}} (-1)^{\mathfrak{m}} w_{x-Q}(\mathfrak{m}) \ , \tag{5.99}$$

where $(-1)^{\mathfrak{m}}$ is the signature of the corresponding permutation. For example:

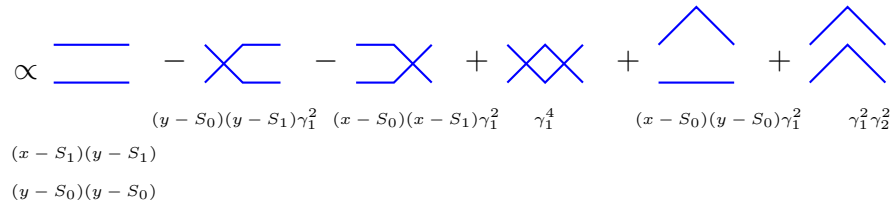
$p_3(x) =$ 

Motzkin paths in $\mathcal{M}^{(1)}_{\mathfrak{S}_n}$ realize only the permutations that are products of commuting transpositions of neighbouring elements of $\{0,\dots,n-1\}$.

The determinant that appears in the expression (5.60) of the self-reproducing kernel $K_{n+1}(x,y)$ is a sum over the set $\mathcal{M}^{(2)}_{\mathfrak{S}_n}$ of disconnected non-intersecting paths with n components of length 2 that correspond to permutations of $\{0,\dots,n-1\}$,

$$\det_{n\times n \text{ submatrix}} (x-Q)(y-Q) = \sum_{\mathfrak{m}\in\mathcal{M}^{(2)}_{\mathfrak{S}_n}} (-1)^{\mathfrak{m}} w_{x-Q,y-Q}(\mathfrak{m}) \ , \tag{5.100}$$

where the notation $w_{x-Q,y-Q}$ means that the first edge of a component has weight $w_{x-Q}(e)$, and the second edge has weight $w_{y-Q}(e)$. For example:

$K_3(x,y) \propto$ 

Other functions of the matrix Q can also be expressed as sums of Motzkin paths. This holds in particular for the elements of the right inverse of $x-Q$, using eq. (5.92).

Continuous fractions

Generating functions of Motzkin paths can be expressed as matrix elements of $\frac{1}{x-Q}$. These matrix elements can themselves be expressed as continuous fractions, for example

$$\left(\frac{1}{x-Q}\right)_{0,0} = \cfrac{1}{x-S_0-\cfrac{\gamma_1^2}{x-S_1-\cfrac{\gamma_2^2}{x-S_2-\cfrac{\gamma_3^2}{x-\dots}}}} \ , \tag{5.101}$$

which results from recursively using the relation

$$M = \left(\begin{array}{c|ccccc} a & b & 0 & \cdots & 0 \\ \hline b & & & & \\ 0 & & & & \\ \vdots & & & N & \\ 0 & & & & \end{array}\right) \quad\Rightarrow\quad (M^{-1})_{0,0} = \frac{1}{a-b^2(N^{-1})_{0,0}} \ , \tag{5.102}$$

starting with $M = x - Q$. Continuous fractions are limits of rational fractions, and the rational fractions that are relevant in our case are minors of $x - Q$,

$$\left(\frac{1}{x-Q}\right)_{0,0} = \lim_{n\to\infty} \frac{\det(x-Q)_{[1,n-1]^2}}{\det(x-Q)_{[0,n-1]^2}} = \lim_{n\to\infty} \frac{\tilde{p}_n(x)}{p_n(x)} \ . \tag{5.103}$$

The polynomials $\tilde{p}_n(x)$ that we just introduced obey the same recursion relation as the orthogonal polynomials $p_n(x)$, with however the initial conditions

$$(\tilde{p}_0, \tilde{p}_1) = (0, 1) \ , \tag{5.104}$$

instead of $(p_{-1}, p_0) = (0, 1)$. Let us also derive an expression of $\left(\frac{1}{x-Q}\right)_{n,n}$ as a continuous fraction, starting with (5.93) $\left(\frac{1}{x-Q}\right)_{n,n} = \psi_n \varphi_n$. The Hilbert transform φ_n of ψ_n obeys the same recursion relation as ψ_n, and therefore as $\frac{p_n}{\sqrt{h_n}}$. So there must exist two functions α, β such that $\sqrt{h_n}\varphi_n = \alpha p_n + \beta \tilde{p}_n$. Let us determine these functions by considering the cases $n = 0, 1$. In the case $n = 0$ we simply have $\alpha(x) = \sqrt{h_0}\varphi_0(x) = h_0 e^{\frac{1}{2}V(x)} \left(\frac{1}{x-Q}\right)_{0,0}$. In the case $n = 1$ we have $\beta = \sqrt{h_1}\varphi_1 - \sqrt{h_0}\varphi_0 p_1$. From the definitions of φ_0, φ_1 as Hilbert transforms, we then deduce

$$\beta(x) = e^{\frac{1}{2}V(x)} \int_\gamma \frac{p_1(x') - p_1(x)}{x - x'} e^{-V(x')} dx' = -e^{\frac{1}{2}V(x)} \int_\gamma e^{-V(x')} dx' = -h_0 e^{\frac{1}{2}V(x)} \ . \tag{5.105}$$

This leads to the desired expression of $\left(\frac{1}{x-Q}\right)_{n,n}$ as a continuous fraction,

$$\left(\frac{1}{x-Q}\right)_{n,n} = \frac{h_0}{h_n} p_n(x)^2 \left(\left(\frac{1}{x-Q}\right)_{0,0} - \frac{\tilde{p}_n(x)}{p_n(x)} \right) \ . \tag{5.106}$$

5.2.4 Christoffel–Darboux formulas

Using the projector Π_N on the first N components, the self-reproducing kernel (5.38) can be written as

$$K_N(x, y) = \vec{\psi}(x)^T \Pi_N \vec{\psi}(y) \ . \tag{5.107}$$

We thus have

$$(y - x) K_N(x, y) = \vec{\psi}(x)^T [\Pi_N, Q] \vec{\psi}(y) \ . \tag{5.108}$$

This involves the **Christoffel–Darboux matrix**

$$\boxed{A^N = [\Pi_N, Q]} \ , \tag{5.109}$$

whose only nonzero coefficients are $A^N_{N-1,N} = -A^N_{N,N-1} = \gamma_N$. And we obtain the **Christoffel–Darboux formula** for the self-reproducing kernel,

$$\boxed{K_N(x, y) = \frac{\sum_{j,k=N-1}^N A^N_{j,k} \psi_j(x) \psi_k(y)}{y - x} = \gamma_N \frac{\psi_{N-1}(x)\psi_N(y) - \psi_N(x)\psi_{N-1}(y)}{y - x}} \ , \tag{5.110}$$

which is more convenient than our original formula (5.38) at large N, since it only involves two terms instead of N terms.

Kernels which can be expressed via a Christoffel–Darboux formula, in terms of finitely many functions ψ_k, are called **integrable kernels**. We will soon see how our constructions based on orthogonal polynomials allow us to reformulate matrix models as integrable systems.

Matrix kernel

Our formulas for the kernel K_N suggest other kernels, which are built using the Hilbert transforms $\vec{\varphi}$ in addition to the original orthogonal functions $\vec{\psi}$. We thus define a matrix kernel

$$\hat{K}_N = \begin{pmatrix} \hat{K}^N_{1,1} & \hat{K}^N_{1,2} \\ \hat{K}^N_{2,1} & \hat{K}^N_{2,2} \end{pmatrix} , \tag{5.111}$$

whose elements are

$$\hat{K}^N_{2,1}(x,y) = K_N(x,y) = \vec{\psi}(x)^T \Pi_N \vec{\psi}(y) , \tag{5.112}$$

$$\hat{K}^N_{2,2}(x,y) = \vec{\psi}(x)^T \Pi_N \vec{\varphi}(y) + \frac{e^{\frac{1}{2}(V(y)-V(x))}}{x-y} , \tag{5.113}$$

$$\hat{K}^N_{1,1}(x,y) = -\vec{\varphi}(x)^T \Pi_N \vec{\psi}(y) + \frac{e^{\frac{1}{2}(V(x)-V(y))}}{x-y} = -\hat{K}^N_{2,2}(y,x) , \tag{5.114}$$

$$\hat{K}^N_{1,2}(x,y) = -\vec{\varphi}(x)^T \Pi_N \vec{\varphi}(y) + \int_\gamma \frac{e^{\frac{1}{2}V(x)} e^{\frac{1}{2}V(y)}}{(x-x')(y-x')} e^{-V(x')} dx' . \tag{5.115}$$

Using the recursion relation for ψ and φ, we obtain the matrix Christoffel–Darboux formula

$$\boxed{(x-y)\hat{K}_N(x,y) = \Psi_N(x)^{-1}\Psi_N(y)} , \tag{5.116}$$

where we introduced the matrix

$$\boxed{\Psi_N = \begin{pmatrix} \psi_{N-1} & \varphi_{N-1} \\ \psi_N & \varphi_N \end{pmatrix}} , \tag{5.117}$$

which obeys $\det \Psi_N(x) = -\frac{1}{\gamma_N}$. The matrix Christoffel–Darboux formula is valid in more general cases, and in particular in multi-matrix models where the matrices $\hat{K}_N$ and Ψ_N can be of rank higher than two.

Heine-type formulas for the matrix kernel

The entries of the matrix kernel obey the Heine-type formulas

$$\hat{K}^N_{2,1}(x,y) = K_N(x,y) = \frac{e^{-\frac{V(x)}{2}} e^{-\frac{V(y)}{2}}}{h_{N-1}} \Big\langle \det(x-M) \det(y-M) \Big\rangle_{N-1\times N-1} , \tag{5.118}$$

$$\hat{K}^N_{2,2}(x,y) = -\hat{K}^N_{1,1}(y,x) = \frac{e^{-\frac{V(x)}{2}} e^{\frac{V(y)}{2}}}{(x-y)} \left\langle \frac{\det(x-M)}{\det(y-M)} \right\rangle_{N\times N} , \tag{5.119}$$

$$\hat{K}^N_{1,2}(x,y) = h_{N-1} e^{\frac{V(x)}{2}} e^{\frac{V(y)}{2}} \left\langle \frac{1}{\det(x-M)\det(y-M)} \right\rangle_{N-1\times N-1} . \tag{5.120}$$

To prove these formulas, the idea is to write their right-hand sides in terms of the functions ψ_k and/or φ_k, and to show that the resulting expressions agree with the Christoffel–Darboux formulas for the kernels. Let us do this in the case of the first formula. Since $\det(x-M)$ is a polynomial function of x of degree N, we can write $(y-x)\Big\langle \det(x-M)\det(y-M)\Big\rangle_{N-1\times N-1}$ as a linear combination of the orthogonal polynomials $p_0(x), \dots, p_N(x)$, and the coefficient of $p_N(x)$ is simply $-\Big\langle \det(y-M)\Big\rangle_{N-1\times N-1}$.

Equivalently, there must exist x-independent functions $c_j(y)$ such that

$$(y-x)\frac{e^{-\frac{V(x)}{2}}e^{-\frac{V(y)}{2}}}{h_{N-1}}\Big\langle \det(x-M)\det(y-M)\Big\rangle_{N-1\times N-1} \\ = -\gamma_N\psi_N(x)\psi_{N-1}(y)+\sum_{j=0}^{N-1}c_j(y)\psi_j(x)\ . \quad (5.121)$$

Let us compute $c_j(y)$ as the scalar product of the left-hand side of this formula with $\psi_j(x)$. Writing $\Big\langle \det(x-M)\det(y-M)\Big\rangle_{N-1\times N-1}$ as an integral over the eigenvalues $\lambda_1,\dots,\lambda_{N-1}$ of M and renaming x as λ_N, we find

$$\sqrt{h_j}\mathcal{Z}_N e^{\frac{V(y)}{2}}c_j(y)=\int_{\gamma^N}d\lambda_1\dots d\lambda_N\Delta_N(\lambda)p_j(\lambda_N)\Delta_{N-1}(\lambda)\prod_{i=1}^{N}(y-\lambda_i)e^{-V(\lambda_i)}\ . \quad (5.122)$$

Symmetrizing the integrand over the N integration variables (exactly as in our proof of Heine's formula), we find $c_{j<N-1}=0$ and $c_{N-1}=\gamma_N\psi_N(y)$. This completes the proof of the Heine-type formula for $K_N(x,y)$.

5.3 Associated integrable system

5.3.1 Differential system and Riemann–Hilbert problem

As a consequence of their recursion and derivation relations, the functions ψ_N must obey a size two differential system of the type

$$\begin{pmatrix}\psi'_{N-1}\\ \psi'_N\end{pmatrix}=D_N\begin{pmatrix}\psi_{N-1}\\ \psi_N\end{pmatrix}\ . \quad (5.123)$$

The system is of size two because the recursion relation (5.53) involves three consecutive functions $\psi_{k+1},\psi_k,\psi_{k-1}$, so that ψ_j can be expressed as a linear combination of ψ_N and ψ_{N-1}, with polynomial coefficients of degrees at most $\max(j-N,N-1-j)$. The derivation relation (5.66) then gives ψ'_k as a combination of ψ_j with $|j-k|\le d$, and this leads to a matrix D_N whose coefficients are polynomials of degrees at most d. We will now determine this matrix more explicitly.

Direct determination of the matrix D_N

From the recursion relation, we can in principle deduce an infinite **folding matrix** $F^N(x)$ such that

$$\vec{\psi}(x)=F^N(x)\vec{\psi}(x)\ , \quad (5.124)$$

$$k\notin\{N-1,N\}\Rightarrow F^N_{j,k}(x)=0\ , \quad (5.125)$$

which means

$$\psi_j(x)=\sum_{k=N-1}^{N}F^N_{j,k}(x)\psi_k(x)\ . \quad (5.126)$$

Let us show that the folding matrix can be expressed in terms of the right and left inverses of $(x-Q)$, and of the Christoffel–Darboux matrix A^N, as

$$F^N(x) = (R(x) - L(x))A^N \ . \tag{5.127}$$

The matrix on the right-hand side satisfies eq. (5.125) due to the zeros of A^N, let us check that it also satisfies eq. (5.124):

$$\begin{aligned}
(R(x) - L(x))A^N\vec{\psi} &= (\vec{\psi}\vec{\varphi}^T - \vec{\varphi}\vec{\psi}^T)\Big((x-Q)\Pi_N - \Pi_N(x-Q)\Big)\vec{\psi} \ , && (5.128)\\
&= (\vec{\psi}\vec{\varphi}^T - \vec{\varphi}\vec{\psi}^T)(x-Q)\Pi_N\vec{\psi} \ , && (5.129)\\
&= \vec{\psi}\vec{\varphi}^T(x-Q)\Pi_N\vec{\psi} \ , && (5.130)\\
&= \vec{\psi}\left(\tfrac{1}{\psi_0}, 0, 0, \cdots\right)\Pi_N\vec{\psi} \ , && (5.131)\\
&= \vec{\psi} \ . && (5.132)
\end{aligned}$$

From the derivation relation $\vec{\psi}' = P\vec{\psi}$, we then deduce an expression for $D_N(x)$ as a size two block of an infinite matrix,

$$D_N(x) = \big(PF^N(x)\big)_{[N-1,N]^2} \ . \tag{5.133}$$

Let us insert our expression for $F^N(x)$, and the formula (5.72) for P. Since $V'(Q)_+$ and $L(x)$ are both strictly upper triangular, we have $\big(V'(Q)_+L(x)A^N\big)_{[N-1,N]^2} = 0$, and similarly $\big(V'(Q)_-R(x)A^N\big)_{[N-1,N]^2} = 0$. Using this remark, we compute

$$\begin{aligned}
D_N(x) &= -\frac{1}{2}\Big((V'(Q)_+ - V'(Q)_-)(R(x) - L(x))A^N\Big)_{[N-1,N]^2} \ , && (5.134)\\
&= -\frac{1}{2}\Big((V'(Q)_+ + V'(Q)_-)(R(x) + L(x))A^N\Big)_{[N-1,N]^2} \ , && (5.135)\\
&= -\frac{1}{2}\Big(V'(Q)(R(x) + L(x))A^N\Big)_{[N-1,N]^2} \ . && (5.136)
\end{aligned}$$

Using eq. (5.92), we can compute $\big(R(x)A^N\big)_{[N-1,N]^2}$ and $\big(L(x)A^N\big)_{[N-1,N]^2}$, and we find $\big((R(x) + L(x))A^N\big)_{[N-1,N]^2} = \sigma_3$ where σ_3 is a Pauli matrix. Moreover, in $R(x) + L(x) = 2(x-Q)^{-1} + \vec{\psi}\vec{\varphi}^T - \vec{\varphi}\vec{\psi}^T$, the last two terms do not contribute to $\Big((V'(x) - V'(Q))(R(x) + L(x))A^N\Big)_{[N-1,N]^2}$. We then find [52]

$$\boxed{D_N(x) = -\frac{1}{2}V'(x)\sigma_3 + \left(\frac{V'(x) - V'(Q)}{x-Q}A^N\right)_{[N-1,N]^2}} \ . \tag{5.137}$$

This shows in particular that D_N is traceless. On the other hand, this is not terribly helpful for computing $\det D_N$, which can be more easily obtained by recursion on N using eq. (5.160). We find

$$\boxed{\operatorname{Tr} D_N = 0} \quad , \quad \boxed{\det D_N = P_N - \frac{1}{4}(V')^2} \ , \tag{5.138}$$

where we define the polynomial

$$\boxed{P_N(x) = \operatorname{Tr}\Pi_N\frac{V'(x) - V'(Q)}{x-Q} = \left\langle \operatorname{Tr}\frac{V'(x) - V'(M)}{x-M}\right\rangle_{N\times N \text{ matrices}}} \ . \tag{5.139}$$

(The last expression comes from the equation (5.56).) The **spectral curve** of the system is now defined by the vanishing of the characteristic polynomial of D_N, and has the equation

$$\boxed{\det(y - D_N(x)) = 0} \,. \tag{5.140}$$

In the appropriate large N limits, this agrees with the spectral curves which we defined in the contexts of the saddle point approximation and of the loop equations, and our polynomial $P_N(x)$ agrees with the previously defined $P_0(x)$ (4.24).

Isomonodromy

For $N > d$, since φ_N satisfies the same derivation relations as ψ_N, the matrix Ψ_N (5.117) obeys the differential equation

$$\boxed{\Psi_N' = D_N \Psi_N} \,. \tag{5.141}$$

Since φ_N is discontinuous (5.86) across the path γ_i, the matrix Ψ_N is discontinuous too, and its behaviour is described by the **monodromy matrix** S_{γ_i},

$$\Psi_N(x + i0) = \Psi_N(x - i0)\, S_{\gamma_i} \quad \text{with} \quad S_{\gamma_i} = \begin{pmatrix} 1 & -2\pi i c_i \\ 0 & 1 \end{pmatrix} \,. \tag{5.142}$$

Our matrix differential equation is **isomonodromic**: the monodromy matrix S_{γ_i} is constant, equivalently D_N is smooth across the path γ_i,

$$\forall x \in \gamma_i \quad D_N(x + i0) = D_N(x - i0) \qquad \Leftrightarrow \qquad S_{\gamma_i}' = 0 \,. \tag{5.143}$$

This equivalence indeed follows from the calculation

$$\begin{aligned} D_N(x + i0) - D_N(x - i0) &= \left((\Psi_N S_{\gamma_i})' (\Psi_N S_{\gamma_i})^{-1} - \Psi_N' \Psi_N^{-1} \right)(x - i0) \,, & (5.144) \\ &= \left(\Psi_N S_{\gamma_i}' S_{\gamma_i}^{-1} \Psi_N^{-1} \right)(x - i0) \,. & (5.145) \end{aligned}$$

Riemann–Hilbert problem

The determination of Ψ_N can also be formulated as a **Riemann–Hilbert problem**. This means that instead of solving the matrix differential equation (5.141), whose coefficient is not necessarily easy to determine, we constrain Ψ_N by its analyticity, monodromy and asymptotic properties:

- Ψ_N is analytic on $\mathbb{C} - \cup_i \gamma_i$, and is everywhere invertible $\det \Psi_N(x) \neq 0$,
- the behaviour of Ψ_N across the path γ_i is given by eq. (5.142),
- the asymptotic behaviour of Ψ_N near infinity is given by

$$\boxed{\Psi_N(x) = \hat{\Psi}_N(x) e^{T_N(x)} \quad \text{with} \quad T_N(x) = \frac{1}{2} \left(2N \log x - V(x) \right) \sigma_3} \,, \tag{5.146}$$

  where $\hat{\Psi}_N(x)$ is analytic near $x = \infty$:

$$\hat{\Psi}_N(x) \sim \begin{pmatrix} 0 & \sqrt{h_{N-1}} \\ \frac{1}{\sqrt{h_N}} & 0 \end{pmatrix} + O\left(\frac{1}{x}\right) \,. \tag{5.147}$$

It can be proved that when it exists, the solution of the Riemann–Hilbert is unique.

The formulation of the conditions on Ψ_N as a Riemann–Hilbert problem is particularly convenient for studying its large N limit. For an interesting and well-defined large N limit to exist, we rescale the potential as $V \to NV$, thereby restoring the same normalization as in the previous Chapters, and simplifying the N-dependence of the asymptotic condition (5.146). Then two functions Ψ_N and $\tilde{\Psi}_N$ which solve two Riemann–Hilbert problems with monodromies and asymptotics that agree up to $(1+O(\frac{1}{N}))$ factors, must themselves be equal up to $(1+O(\frac{1}{N}))$ factors. The difficulty is to guess a correct and manageable function $\tilde{\Psi}_N$, which solves an approximate problem. To do that, we can use non-rigorous methods such as the saddle point approximation, or formal solutions of loop equations. This then provides a rigorous solution to the determination of the large N behaviour of Ψ_N. This Riemann–Hilbert asymptotic method has been a major tool for random matrix asymptotic analysis, under the name of the Deift–Zhou steepest descent method [53].

5.3.2 Deformations and flat connections

We will now study how the matrix Ψ_N depends on the coupling constants t_k in the potential $V(x) = \sum_k \frac{t_k}{k} x^k$, and on the matrix size N. We will see that t_k and N play the roles of time variables in the integrable system – a discrete time variable in the case of N.

Dependence on the coupling constants t_k

In the same way as Ψ_N obeys the matrix differential equation (5.141) in the variable x, it obeys a differential equation in t_k,

$$\boxed{\frac{\partial}{\partial t_k}\Psi_N = H_{N,k}\Psi_N} \ . \tag{5.148}$$

The existence of such an equation is trivial, as we can set $H_{N,k} = \frac{\partial}{\partial t_k}\Psi_N \Psi_N^{-1}$. The equation will become interesting after we determine $H_{N,k}$ and find that it is a polynomial function of x of degree k. Let us start by computing $\frac{\partial \psi_N}{\partial t_k}$. Since $\psi_N \propto p_N e^{-\frac{1}{2}V}$ (5.39) where $\partial_{t_k} p_N(x)$ is a polynomial of degree $N-1$ and $\partial_{t_k} V = \frac{1}{k}x^k$, there is an infinite matrix $\mathcal{H}_k$ such that

$$\frac{\partial \psi_N}{\partial t_k} = \sum_{j=0}^{N+k} (\mathcal{H}_k)_{N,j}\ \psi_j \ , \tag{5.149}$$

and $\mathcal{H}_k + \frac{Q^k}{2k}$ must be strictly lower triangular. Using $\partial_{t_k}\int_\gamma \psi_i\psi_j = 0$, we moreover see that $\mathcal{H}_k$ must be antisymmetric,

$$\mathcal{H}_k = -\mathcal{H}_k^T \ . \tag{5.150}$$

This implies

$$\mathcal{H}_k = -\frac{1}{2k}\left((Q^k)_+ - (Q^k)_-\right) \ . \tag{5.151}$$

As in our computation of D_N, we obtain $H_{N,k}$ by multiplying with the folding matrix,

$$H_{N,k}(x) = \left(\mathcal{H}_k F^N(x)\right)_{[N-1,N]^2} = \left(\mathcal{H}_k(R(x)-L(x))A^N\right)_{[N-1,N]^2} \ , \tag{5.152}$$

and we find

$$\boxed{H_{N,k}(x) = -\frac{x^k}{2k}\sigma_3 + \left(\frac{x^k - Q^k}{x - Q}A^N\right)_{[N-1,N]^2}} \ . \tag{5.153}$$

The coupling constant t_k parametrize **isomonodromic deformations** of our isomonodromic differential equation: $H_{N,k}(x)$ behaves smoothly when x crosses γ_i, equivalently the monodromy matrix S_{γ_i} of Ψ_N does not depend on t_k. (This equivalence is proved in the same way as the analogous equivalence (5.143) for $D_N(x)$.)

The matrix coefficients $H_{N,k}$ are moreover constrained by the compatibility of the equations (5.148) and (5.141) that Ψ_N obeys. Namely, two differential operators which kill Ψ_N must commute with each other, and we must have

$$\left[\frac{\partial}{\partial t_k} - H_{N,k}, \frac{\partial}{\partial t_j} - H_{N,j}\right] = \frac{\partial}{\partial t_j}H_{N,k} - \frac{\partial}{\partial t_k}H_{N,j} - [H_{N,j}, H_{N,k}] = 0 \ , \tag{5.154}$$

$$\left[\frac{\partial}{\partial x} - D_N, \frac{\partial}{\partial t_k} - H_{N,k}\right] = \frac{\partial}{\partial t_k}D_N - \frac{\partial}{\partial x}H_{N,k} + [D_N, H_{N,k}] = 0 \ . \tag{5.155}$$

This can be reformulated as the flatness of the following connection of a $GL_2(\mathbb{C})$ bundle over the space $\mathbb{C}^{d+1}$ with coordinates x and t_k's,

$$\nabla = d - D_N dx - \sum_k H_{N,k} dt_k \ . \tag{5.156}$$

Introducing the operator

$$L = \frac{\partial}{\partial x} - D_N \ , \tag{5.157}$$

the second compatibility conditions can also be reformulated as the **Lax equations**

$$\frac{\partial}{\partial t_k}L = [H_{N,k}, L] \ , \tag{5.158}$$

where D_N plays the role of the **Lax matrix**. Equivalently, the connection $d - D_N dx$ can be interpreted as the Higgs field in the **Hitchin system** of the group $GL_2(\mathbb{C})$ over the complex plane $\mathbb{C}$. In this context, the times t_k parametrize the moduli space of connections.

Dependence on the discrete time N

From the recursion relations (5.53) and (5.82) for ψ_N and φ_N, we deduce the recursion relation for Ψ_N,

$$\boxed{\Psi_{N+1} = R_N \Psi_N \quad \text{with} \quad R_N = \begin{pmatrix} 0 & 1 \\ -\frac{\gamma_N}{\gamma_{N+1}} & \frac{x - S_N}{\gamma_{N+1}} \end{pmatrix}} \ . \tag{5.159}$$

The matrix R_N is now a polynomial of x of degree one, and plays the same role for the discrete time N as the matrices D_N and $H_{N,k}$ for the variables x and t_k respectively. The compatibility of the discrete time evolution with the dependences on x and t_k leads to difference equations for D_N and $H_{N,k}$,

$$D_{N+1} = \partial_x R_N \, R_N^{-1} + R_N D_N R_N^{-1} \ , \tag{5.160}$$

$$H_{N+1,k} = \partial_{t_k} R_N \, R_N^{-1} + R_N H_{N,k} R_N^{-1} \ . \tag{5.161}$$

These equations can be interpreted as flatness conditions involving differences in addition to differentials.

5.3.3 Tau function and Baker–Akhiezer function

Tau function

Given an isomonodromic matrix differential equation (5.141) with an analytic matrix coefficient D_N, and a solution Ψ_N whose asymptotic behaviour at infinity is governed by a function T_N as in eq. (5.146), there is a **Tau function** τ_N such that

$$\boxed{\frac{\partial}{\partial t_k}\log\tau_N = \frac{1}{2\pi i}\oint_\infty dx\ \mathrm{Tr}\,\Psi_N^{-1}(x)\Psi_N'(x)\frac{\partial}{\partial t_k}T_N(x)}\ , \tag{5.162}$$

for any isomonodromic deformation $\frac{\partial}{\partial t_k}$. The existence of the Tau function was established by Jimbo–Miwa [54] and Ueno–Takasaki [55], by showing that the formula for $\frac{\partial}{\partial t_k}\log\tau_N$ is compatible with the identity $\partial_{t_k}\partial_{t_j}\log\tau_N = \partial_{t_j}\partial_{t_k}\log\tau_N$. Tau functions are fundamental objects in integrable systems, analogous to partition functions in statistical physics [56]. And we will now show that the Tau function of our matrix differential equation is closely related to the partition function of our matrix model.

The function T_N obeys the identity $\frac{\partial}{\partial t_k}T_N(x) = \frac{x^k}{kV'(x)}T_N'(x)$, which implies

$$\frac{\partial}{\partial t_k}\log\tau_N = \frac{1}{2\pi i}\oint_\infty dx\,\frac{x^k}{kV'(x)}\,\mathrm{Tr}\,\hat\Psi_N^{-1}\hat\Psi_N'T_N'(x)\ . \tag{5.163}$$

The trace which appears in the last expression can now be identified with the last term in

$$\mathrm{Tr}\,D_N^2 = \mathrm{Tr}\left(\Psi_N'\Psi_N^{-1}\right)^2 = \mathrm{Tr}\left(\left(\hat\Psi_N'\hat\Psi_N^{-1}\right)^2 + (T_N')^2 + 2\hat\Psi_N^{-1}\hat\Psi_N'T_N'\right)\ . \tag{5.164}$$

However, the first two terms do not contribute to our residue at $x=\infty$ – in particular, since $\hat\Psi(x)$ is analytic at $x=\infty$, we have $\hat\Psi_N^{-1}\hat\Psi_N'(x) \underset{x\to\infty}{=} O(\frac{1}{x^2})$. We thus obtain

$$\frac{\partial}{\partial t_k}\log\tau_N = \frac{1}{2\pi i}\oint_\infty dx\,\frac{x^k}{2kV'(x)}\,\mathrm{Tr}\,D_N(x)^2\ . \tag{5.165}$$

But $\mathrm{Tr}\,D_N^2 = -2\det D_N$ was computed in eq. (5.138), which implies

$$\frac{\partial}{\partial t_k}\log\tau_N = \frac{1}{2\pi i}\oint_\infty dx\,\frac{x^k}{kV'(x)}\left(\frac{V'(x)^2}{4} - \left\langle\mathrm{Tr}\,\frac{V'(x)-V'(M)}{x-M}\right\rangle_{N\times N}\right)\ , \tag{5.166}$$

$$= -\frac{1}{k}\left\langle\mathrm{Tr}\,M^k\right\rangle_{N\times N} + \frac{1}{2\pi i}\oint_\infty dx\,\frac{x^k}{kV'(x)}\left\langle\mathrm{Tr}\,\frac{V'(M)}{x-M}\right\rangle_{N\times N}\ . \tag{5.167}$$

The remaining residue at infinity most often vanishes as a result of $k \leq d+1$ and $V(x) \underset{x\to\infty}{=} t_{d+1}x^{d+1} + O(x^d)$. To compute it, let us write $\frac{V'(M)}{x-M} = \frac{V'(M)}{x} + \frac{MV'(M)}{x^2} + O\left(\frac{1}{x^3}\right)$. Using the equations (5.56) and (5.74), we find

$$\left\langle\mathrm{Tr}\,V'(M)\right\rangle_{N\times N} = \mathrm{Tr}\,V'(Q)\Pi_N = 0\ , \tag{5.168}$$

$$\left\langle\mathrm{Tr}\,MV'(M)\right\rangle_{N\times N} = \mathrm{Tr}\,QV'(Q)\Pi_N = \sum_{k=0}^{N-1}\left(\gamma_{k+1}\frac{k+1}{\gamma_{k+1}} + \gamma_k\frac{k}{\gamma_k}\right) = N^2\ , \tag{5.169}$$

and therefore

$$\frac{\partial}{\partial t_k}\log\tau_N = -\frac{1}{k}\left\langle\mathrm{Tr}\,M^k\right\rangle_{N\times N} + \frac{N^2\delta_{k,d+1}}{(d+1)t_{d+1}}\ . \tag{5.170}$$

This shows that the Tau function is closely related to the partition function $\mathcal{Z}_N$,

$$\boxed{t_{d+1}^{-\frac{N^2}{d+1}} \tau_N \propto \mathcal{Z}_N = \int_{H_N(\gamma)} dM\, e^{-\operatorname{Tr} V(M)}} \,, \tag{5.171}$$

where the proportionality coefficient is a function of N and of the path γ.

Baker–Akhiezer function

Let us express the orthogonal polynomials in terms of the Tau function, in order to find their interpretation in the associated isomonodromic integrable system. According to Heine's formula, we have

$$p_N(x) = \Big\langle \det(x - M) \Big\rangle_{N\times N} = \frac{1}{\mathcal{Z}_N} \int_{H_N(\gamma)} dM\, e^{\operatorname{Tr}\log(x-M) - \operatorname{Tr} V(M)} \,, \tag{5.172}$$

$$= \frac{1}{\mathcal{Z}_N} \int_{H_N(\gamma)} dM\, e^{N\log x - \operatorname{Tr}\sum_{k=1}^{\infty} \left(t_k - \frac{1}{x^k}\right)\frac{M^k}{k}} \,. \tag{5.173}$$

This matrix integral can be interpreted as the value of our partition function, after performing an x-dependent shift of the coefficients t_k of the potential. In terms of the Tau function $\tau_N(\vec{t} = (t_1, t_2, \cdots))$, we therefore have

$$\boxed{p_N(x) = x^N \frac{\tau_N(\vec{t} + [x])}{\tau_N(\vec{t})}} \,, \tag{5.174}$$

where we introduced Sato's notation

$$[x] = \left(-\frac{1}{x}, -\frac{1}{x^2}, -\frac{1}{x^3}, \dots\right) \,. \tag{5.175}$$

This formula for $p_N(x)$ coincides with the **Sato formula** which expresses the **Baker–Akhiezer function** of an integrable system in terms of its Tau function. Therefore, the orthogonal polynomial $p_N(x)$ is the Baker–Akhiezer function of the associated integrable system. Similarly, the Heine-type formula for the Hilbert transform φ_{N-1} leads to

$$\varphi_{N-1}(x) \frac{e^{-\frac{1}{2}V(x)}}{\sqrt{h_{N-1}}} = x^{-N} \frac{\tau_N(\vec{t} - [x])}{\tau_N(\vec{t})} \,, \tag{5.176}$$

which coincides with the Sato formula for the **dual Baker–Akhiezer function**. And Heine-type formulas for the matrix kernel allow it to be expressed in terms of the Tau function. For example, eq. (5.119) for the matrix kernel element $\hat{K}^N_{1,1}$ leads to

$$e^{\frac{1}{2}V(x)} e^{-\frac{1}{2}V(y)} \hat{K}^N_{1,1}(y,x) = \frac{1}{y-x} \frac{\tau_N(\vec{t} + [x] - [y])}{\tau_N(\vec{t})} \,. \tag{5.177}$$

Hirota equation

Let us see how the orthogonality of orthogonal polynomial leads to powerful equations for Tau functions. Let $\vec{t}$ and $\vec{\tilde{t}}$ be two families of times, with $V(x)$ and $\tilde{V}(x)$ the corresponding potentials, and p_k and $\tilde{p}_k$ the corresponding families of orthogonal polynomials. We have

$$\forall m \leq n, \quad h_n \delta_{n,m} = \int_\gamma dx'\, p_n(x') \tilde{p}_m(x') e^{-V(x')} \,, \tag{5.178}$$

$$= \int_\gamma dx' \operatorname*{Res}_{x=x'} \frac{dx}{x - x'} e^{\tilde{V}(x) - V(x)} p_n(x) \tilde{p}_m(x') e^{-\tilde{V}(x')} \,, \tag{5.179}$$

$$= -\operatorname*{Res}_{x=0} dx\, e^{\tilde{V}(x) - V(x)} p_n(x) \tilde{h}_m^{-\frac{1}{2}} e^{-\frac{1}{2}\tilde{V}(x)} \tilde{\varphi}_m(x) \,. \tag{5.180}$$

We have been able to exchange the integral over γ and the residue, and to obtain a residue at $x = 0$, by treating the integrand as a formal series in $\frac{1}{x}$. Using the expressions of p_n and φ_m in terms of Tau functions, we obtain

$$\forall m \leq n, \quad h_n \delta_{n,m} \tau_n(\vec{t}) \tau_{m+1}(\vec{\tilde{t}}) = -\operatorname*{Res}_{x=0} \frac{dx}{x^{m+1-n}} e^{\tilde{V}(x)-V(x)} \tau_n(\vec{t} + [x]) \tau_{m+1}(\vec{\tilde{t}} - [x]) \ . \tag{5.181}$$

Renaming the times as $(\vec{t}, \vec{\tilde{t}}) \to (\vec{t} - \vec{u}, \vec{t} + \vec{u})$, we obtain the **Hirota equation**

$$\boxed{\begin{aligned} \forall m \leq n, \qquad & h_n \delta_{n,m} \tau_n(\vec{t} - \vec{u}) \tau_{m+1}(\vec{t} + \vec{u}) \\ & = -\operatorname*{Res}_{x=0} \frac{dx}{x^{m+1-n}} e^{2\sum_k \frac{u_k x^k}{k}} e^{\sum_k x^{-k} \frac{\partial}{\partial u_k}} \tau_n(\vec{t} - \vec{u}) \tau_{m+1}(\vec{t} + \vec{u}) \ . \end{aligned}} \tag{5.182}$$

Hirota equations, sometimes called bilinear equations, are ubiquitous in integrable systems, to the extent that integrability can be defined by the existence of a Hirota equation. Our particular one-matrix model, and its size two matrix differential equation, has one family of times $\vec{t}$, and our Hirota equation only involves residues at $x = 0$. Expanding our Hirota equation in powers of the u_ks yields an infinite hierarchy of quadratic partial differential equations for the Tau functions, called the **KdV hierarchy**. More general integrable systems have Hirota equations that involve singularities at other values of x, and that encode other hierarchies of partial differential equations [56].

Determinantal formulas

In the language of Tau functions, the Christoffel–Darboux formula translates into

$$\tau_N(\vec{t}) \tau_N(\vec{t} + [x] - [y]) = \tau_{N-1}(\vec{t} + [x]) \tau_{N+1}(\vec{t} - [y]) - \tau_N(\vec{t} + [x]) \tau_N(\vec{t} - [y]) \ . \tag{5.183}$$

Moreover, the identity $\hat{K}_{1,1}^{N+1}(x,y) - \hat{K}_{1,1}^{N}(x,y) = \psi_N(x)\varphi_N(y)$ (which follows from the definition of the matrix kernel) translates into

$$\tau_N(\vec{t}) \tau_{N+1}(\vec{t} + [x] - [y]) - \tau_{N+1}(\vec{t}) \tau_N(\vec{t} + [x] - [y]) = (y - x) \tau_N(\vec{t} + [x]) \tau_{N+1}(\vec{t} - [y]) \ . \tag{5.184}$$

After some tedious algebra, it is possible to eliminate τ_{N-1} and τ_{N+1}, and to obtain a quadratic equation for the Tau function τ_N,

$$\begin{aligned} & \frac{(x - x')(y - y')}{(x - y)(x - y')(x' - y)(x' - y')} \tau_N(\vec{t}) \tau_N(\vec{t} + [x] + [x'] - [y] - [y']) \\ & \quad = \frac{\tau_N(\vec{t} + [x'] - [y]) \tau_N(\vec{t} + [x] - [y'])}{(x' - y)(x - y')} - \frac{\tau_N(\vec{t} + [x] - [y]) \tau_N(\vec{t} + [x'] - [y'])}{(x - y)(x' - y')} \ . \end{aligned} \tag{5.185}$$

We call this a determinantal formula, as the right-hand side is the determinant of a matrix of size two. Actually, iterating this formula leads to the more general **determinantal formula**

$$\boxed{\frac{\tau_N(\vec{t} + \sum_{i=1}^k [x_i] - \sum_{i=1}^k [y_i])}{\tau_N(\vec{t})} = \frac{\det_{1 \leq i,j \leq k} \left(\frac{\tau_N(\vec{t} + [x_i] - [y_j])}{(x_i - y_j)} \right)}{\det_{1 \leq i,j \leq k} \left(\frac{1}{x_i - y_j} \right)}} \ , \tag{5.186}$$

whose right-hand side involves the **Cauchy determinant**,

$$\det_{1\le i,j\le k}\left(\frac{1}{x_i-y_j}\right)=\prod_{i<j}^{k}(x_i-x_j)(y_i-y_j)\prod_{i,j}^{k}(x_i-y_j)^{-1}\,. \tag{5.187}$$

The determinantal formula for the Tau function is a generalization of Fay's trisecant identity for theta functions. The theta functions of Fay's identity are indeed Tau functions of isospectral integrable systems [56].

5.3.4 Correlation functions

Determinantal formulas

We will now describe how the connected (4.21) and disconnected (4.22) correlation functions are related to the observables of our integrable system. In the matrix integral, the correlation functions involve insertions of $\mathrm{Tr}\,\frac{1}{x-M}$, whereas Heine-type formulas involve insertions of $\det(x-M)$. These quantities are related by

$$\mathrm{Tr}\,\frac{1}{x-M}=\lim_{y\to x}\frac{1}{x-y}\left(\frac{\det(x-M)}{\det(y-M)}-1\right)\,. \tag{5.188}$$

So, using the Heine-type formula (5.119), we have

$$W_1(x)=\lim_{y\to x}\left(e^{\frac{1}{2}V(x)-\frac{1}{2}V(y)}\hat{K}^N_{1,1}(y,x)-\frac{1}{x-y}\right)\,. \tag{5.189}$$

Using the expression (5.116) for the matrix kernel, we compute the limit and find

$$W_1(x)=\frac{V'(x)}{2}+\left(\Psi_N^{-1}(x)\Psi_N'(x)\right)_{1,1}\,. \tag{5.190}$$

Using the matrix differential equation (5.141) for Ψ_N, this leads to

$$\boxed{W_1(x)=\frac{V'(x)}{2}+\mathrm{Tr}\,D_N(x)M_N(x)}\,, \tag{5.191}$$

where we introduce the projector

$$\boxed{M_N=\Psi_N\begin{pmatrix}1&0\\0&0\end{pmatrix}\Psi_N^{-1}}\quad\Rightarrow\quad\begin{cases}M_N^2=M_N\,,\\ \mathrm{Tr}\,M_N=1\,,\\ \det M_N=0\,,\end{cases}\quad\text{and}\quad M_N'=[D_N,M_N]\,. \tag{5.192}$$

Similarly, let us compute the disconnected two-point function, starting with its expression in terms of insertions of determinants,

$$\hat{W}_2(x,x')=\lim_{\substack{y\to x\\ y'\to x'}}\frac{1}{(x-y)(x'-y')}\left\langle\left(\frac{\det(x-M)}{\det(y-M)}-1\right)\left(\frac{\det(x'-M)}{\det(y'-M)}-1\right)\right\rangle\,. \tag{5.193}$$

Writing this in terms of Tau functions, and using the quadratic equation (5.185), we obtain

$$\begin{aligned}\hat{W}_2(x,x')=\lim_{\substack{y\to x\\ y'\to x'}}\Bigg\{&\frac{(x-y')(x'-y)}{(x-x')(y'-y)}e^{\frac{V(x)+V(x')}{2}-\frac{V(y)+V(y')}{2}}\\ &\times\left(\hat{K}^N_{1,1}(x,y)\hat{K}^N_{1,1}(x',y')-\hat{K}^N_{1,1}(x,y')\hat{K}^N_{1,1}(x',y)\right)\\ &-e^{\frac{1}{2}V(x)-\frac{1}{2}V(y)}\frac{\hat{K}^N_{1,1}(x,y)}{(x'-y')}-e^{\frac{1}{2}V(x')-\frac{1}{2}V(y')}\frac{\hat{K}^N_{1,1}(x,y)}{(x'-y')}+\frac{1}{(x-y)(x'-y')}\Bigg\}\,,\end{aligned} \tag{5.194}$$

Performing the limits, we obtain

$$\hat{W}_2(x,x') = W_1(x)W_1(x') - \hat{K}^N_{1,1}(x,x')\hat{K}^N_{1,1}(x',x) - \frac{1}{(x-x')^2} \ . \tag{5.195}$$

Equivalently, the connected two-point function is $W_2(x,x') = -\hat{K}^N_{1,1}(x,x')\hat{K}^N_{1,1}(x',x) - \frac{1}{(x-x')^2}$, which can be rewritten in terms of the projector M_N as

$$\boxed{W_2(x,x') = \frac{\operatorname{Tr} M_N(x)M_N(x')}{(x-x')^2} - \frac{1}{(x-x')^2}} \ . \tag{5.196}$$

Similarly, writing the correlation function $W_{n\geq 3}$ in terms of insertions of determinants, using the determinantal formula (5.186) for the corresponding Tau functions, and writing the result in terms of the projector M_N, we would obtain the **determinantal formula**

$$\boxed{W_{n\geq 3}(x_1,\cdots x_n) = \sum_{\sigma\in\mathfrak{S}_n} (-1)^\sigma \frac{\operatorname{Tr}\prod_{i=1}^n M_N(x_{\sigma(i)})}{\prod_{i=1}^n (x_{\sigma(i)} - x_{\sigma(i+1)})}} \ . \tag{5.197}$$

Loop equations

Our formulas for the correlation functions in terms of the projector M_N are compatible with the loop equations. Let us check this in the case of the first loop equation 4.26. From eq. (5.196) and the properties of M_N, we see that $W_2(x,x')$ has no pole at $x=x'$, and compute

$$W_2(x,x) = \frac{1}{2}\operatorname{Tr} M_N(x)M_N''(x) = \operatorname{Tr}\left(M_N^2 D_N^2 - (M_N D_N)^2\right)(x) \ . \tag{5.198}$$

Now, using the characteristic equation of size two matrices $A^2 = A\operatorname{Tr} A - \mathrm{Id}\det A$, we find $\operatorname{Tr} M_N^2 D_N^2 = -\det D_N$ and $\operatorname{Tr}(M_N D_N)^2 = (\operatorname{Tr} M_N D_N)^2$, and therefore

$$W_2(x,x) = -\det D_N(x) - \left(W_1(x) - \frac{V'(x)}{2}\right)^2 \ . \tag{5.199}$$

Using the formula (5.138) for $\det D_N(x)$, this can be rewritten as

$$W_1(x)^2 + W_2(x,x) = V'(x)W_1(x) - P_N(x) \ . \tag{5.200}$$

This agrees with the first loop equation, provided the different normalization of the potential V is taken into account.

Asymptotic expansion of the matrix kernel

Conversely, we could consider the correlation functions as known, and deduce the matrix kernel. To do this, we start with the identity

$$\frac{\det(x-M)}{\det(y-M)} = \exp\int_y^x \operatorname{Tr}\frac{1}{x'-M}dx' \ . \tag{5.201}$$

Therefore, the Heine-type formula (5.119) for the matrix kernel element $\hat{K}^N_{2,2}$ leads to

$$(x-y)e^{\frac{V(x)}{2}}e^{-\frac{V(y)}{2}}\hat{K}^N_{2,2}(x,y) = 1 + \sum_{n=1}^\infty \frac{1}{n!}\left(\prod_{i=1}^n \int_y^x dx_i\right)\hat{W}_n(x_1,\cdots x_n) \ , \tag{5.202}$$

$$= \exp\left\{\sum_{n=1}^\infty \frac{1}{n!}\left(\prod_{i=1}^n \int_y^x dx_i\right) W_n(x_1,\cdots x_n)\right\} \ . \tag{5.203}$$

This is in particular useful for computing the large N asymptotic expansion of the matrix kernel, using the topological expansion of correlation functions. In the case of $\hat{K}^N_{2,2}$, we obtain

$$(x-y)e^{\frac{V(x)}{2}}e^{-\frac{V(y)}{2}}\hat{K}^N_{2,2}(x,y) \\ = \exp\left\{\sum_{g=0}^{\infty}\sum_{n=1}^{\infty}\frac{N^{2-2g-n}}{n!}\left(\prod_{i=1}^{n}\int_y^x dx_i\right)W_{g,n}(x_1,\cdots x_n)\right\} . \quad (5.204)$$

This is a type of WKB expansion, where the matrix size N plays the role of the Planck constant.

5.4 Multi-matrix models

In a multi-matrix models, the partition function has an eigenvalue representation of the type

$$\boxed{\mathcal{Z}_N = \frac{1}{(N!)^2}\int_{\gamma^N\times\tilde{\gamma}^N} d\lambda_0\cdots d\lambda_{N-1}d\tilde{\lambda}_0\cdots d\tilde{\lambda}_{N-1} \\ \Delta(\lambda_0,\cdots\lambda_{N-1})\Delta(\tilde{\lambda}_0,\cdots\tilde{\lambda}_{N-1})\det_{i,j=0,\cdots N-1}\omega(\tilde{\lambda}_j,\lambda_i)\ .} \quad (5.205)$$

In the case of a two-matrix model, this follows from the argument in Section 6.1.3, and the function $\omega(\tilde{\lambda},\lambda)$ is given in terms of the potentials V_1, V_2 by

$$\omega(\tilde{\lambda},\lambda) = e^{-V_1(\lambda)}e^{-V_2(\tilde{\lambda})}e^{\lambda\tilde{\lambda}}\ . \quad (5.206)$$

In the more general case of a matrix chain of length L, this function is

$$\omega(\mu_L,\mu_1) = \int_{\gamma^{L-2}} d\mu_2 d\mu_3\cdots d\mu_{L-1}\prod_{i=1}^{L}e^{-V_i(\mu_i)}\prod_{i=1}^{L-1}e^{\mu_i\mu_{i+1}}\ . \quad (5.207)$$

5.4.1 Determinantal formulas

The determinantal formula (5.11) for the partition function still holds in multi-matrix models, provided the matrix elements $H_{i,j} = \langle\tilde{p}_i|p_j\rangle$ are computed from the scalar product

$$\boxed{\langle f|g\rangle = \int_{\gamma\times\tilde{\gamma}} d\lambda d\tilde{\lambda}\ f(\tilde{\lambda})g(\lambda)\omega(\tilde{\lambda},\lambda) = f\,\tilde{\star}\,\omega\star g}\ . \quad (5.208)$$

However, in the case where p_i and $\tilde{p}_i$ are monomials, this determinant is no longer a Hankel or Toeplitz determinant.

Let us evaluate the joint eigenvalues distributions

$$R_{k,k'}(\lambda_0,\cdots\lambda_{k-1},\tilde{\lambda}_0,\cdots\tilde{\lambda}_{k'-1}) = \frac{1}{\mathcal{Z}_N}\frac{1}{N!^2}\int_{\gamma^{N-k}\times\tilde{\gamma}^{N-k'}} d\lambda_k\cdots d\lambda_{N-1}d\tilde{\lambda}_{k'}\cdots d\tilde{\lambda}_{N-1} \\ \Delta(\lambda_0,\cdots\lambda_{N-1})\Delta(\tilde{\lambda}_0,\cdots\tilde{\lambda}_{N-1})\det_{i,j=0,\ldots N-1}\omega(\lambda_i,\tilde{\lambda}_j)\ . \quad (5.209)$$

We will use not only the self-reproducing kernel K (5.16) which is associated to our two families of monic polynomials p_j and $\tilde{p}_j$, but also the kernels

$$\boxed{\begin{aligned} H &= K\,\tilde{\star}\,\omega\ , \quad (5.210)\\ \tilde{H} &= \omega\star K\ , \quad (5.211)\\ J &= -\omega+\omega\star K\,\tilde{\star}\,\omega\ . \quad (5.212)\end{aligned}}$$

These kernels form a self-reproducing family, in the sense that

$$H \star H = H \ , \tag{5.213}$$
$$H \star K = K \ , \tag{5.214}$$
$$\tilde{H} \tilde{\star} \tilde{H} = \tilde{H} \ , \tag{5.215}$$
$$K \tilde{\star} \tilde{H} = K \ , \tag{5.216}$$
$$J \star H = J \star K = \tilde{H} \tilde{\star} J = K \tilde{\star} J = 0 \ . \tag{5.217}$$

In addition, we have

$$\int d\lambda \ H(\lambda, \lambda) = \int d\tilde{\lambda} \ \tilde{H}(\tilde{\lambda}, \tilde{\lambda}) = N \ . \tag{5.218}$$

Dyson's formula for the joint eigenvalue distributions, and its proof, can be generalized, and we obtain the **Eynard–Mehta theorem** [57, 1],

$$\boxed{R_{k,k'} = \frac{(N-k)!(N-k')!}{N!^2} \det_{\substack{i,j=0,\cdots k-1 \\ i',j'=0,\cdots k'-1}} \begin{pmatrix} H(\lambda_i, \lambda_j) & K(\lambda_i, \tilde{\lambda}_{j'}) \\ J(\tilde{\lambda}_{i'}, \lambda_j) & \tilde{H}(\tilde{\lambda}_{i'}, \tilde{\lambda}_{j'}) \end{pmatrix}} \ , \tag{5.219}$$

which involves the determinant of a matrix of size $k + k'$. As special cases, we have the total joint distribution

$$N!^2 R_{N,N} = \det \begin{pmatrix} H & K \\ J & \tilde{H} \end{pmatrix} = \det \begin{pmatrix} K \tilde{\star} \omega & K \\ -\omega + \tilde{H} \tilde{\star} \omega & \tilde{H} \end{pmatrix} = \det \begin{pmatrix} 0 & K \\ -\omega & \tilde{H} \end{pmatrix} \ , \tag{5.220}$$
$$= \det_{i,j=0,\cdots N-1} K(\lambda_i, \tilde{\lambda}_j) \det_{i,j=0,\cdots N-1} \omega(\lambda_i, \tilde{\lambda}_j) \ , \tag{5.221}$$

and the joint distributions of the λ's and $\tilde{\lambda}$'s, respectively,

$$R_{k,0} = \frac{(N-k)!}{N!} \det_{i,j=0,\cdots k-1} H(\lambda_i, \lambda_j) \ , \tag{5.222}$$
$$R_{0,k} = \frac{(N-k)!}{N!} \det_{i,j=0,\cdots k-1} \tilde{H}(\tilde{\lambda}_i, \tilde{\lambda}_j) \ . \tag{5.223}$$

5.4.2 Biorthogonal polynomials

Let **biorthogonal polynomials** be monic polynomials $p_k, \tilde{p}_k$ such that $\deg p_k = \deg \tilde{p}_k = k$ and

$$\boxed{\langle \tilde{p}_k | p_{k'} \rangle = h_k \delta_{k,k'}} \ . \tag{5.224}$$

As in the case of orthogonal polynomials, the partition function is then a product of the coefficients h_k (5.37). Biorthogonal polynomials exist if and only if $\forall N \ \mathcal{Z}_N \neq 0$, and for a given function $\omega(\tilde{\lambda}, \lambda)$ this condition is obeyed almost everywhere in the space of integration contours. Let us now see how the orthogonal polynomials can be computed.

Generalized Heine's formula

In our multi-matrix model, let us denote the expectation value of a function of $\lambda = (\lambda_0, \cdots \lambda_{N-1})$ and $\tilde{\lambda} = (\tilde{\lambda}_0, \cdots \tilde{\lambda}_{N-1})$ as

$$\left\langle f(\lambda, \tilde{\lambda}) \right\rangle_N = \frac{1}{\mathcal{Z}_N} \frac{1}{(N!)^2} \int_{\gamma^N \times \tilde{\gamma}^N} d\lambda d\tilde{\lambda} \ \Delta(\lambda) \Delta(\tilde{\lambda}) \det_{i,j=0,\cdots N-1} \omega(\tilde{\lambda}_j, \lambda_i) \times f(\lambda, \tilde{\lambda}) \ . \tag{5.225}$$

The reasoning that leads to Heine's formula can be generalized to multi-matrix models, and leads to the **generalized Heine's formula**,

$$p_N(x) = \left\langle \prod_{i=0}^{N-1} (x - \lambda_i) \right\rangle_N \quad , \tag{5.226}$$

$$\tilde{p}_N(y) = \left\langle \prod_{i=0}^{N-1} (y - \tilde{\lambda}_i) \right\rangle_N \quad . \tag{5.227}$$

We also have determinantal formulas involving the moments $M_{k,l} = \left\langle \lambda_1^k \mu_1^l \right\rangle_N$,

$$p_k(x) = \frac{\det \begin{pmatrix} M_{0,0} & M_{0,1} & M_{0,2} & \dots & M_{0,k-1} & 1 \\ M_{1,0} & M_{1,1} & & & M_{1,k-1} & x \\ M_{2,0} & & & & M_{2,k-1} & x^2 \\ \vdots & & & & \vdots & \vdots \\ M_{k,0} & & & & M_{k,k-1} & x^k \end{pmatrix}}{\det \begin{pmatrix} M_{0,0} & M_{0,1} & M_{0,2} & \dots & M_{0,k-1} \\ M_{1,0} & M_{1,1} & & & M_{1,k-1} \\ M_{2,0} & & & & M_{2,k-1} \\ \vdots & & & & \vdots \\ M_{k-1,0} & & & & M_{k-1,k-1} \end{pmatrix}} \quad , \tag{5.228}$$

and a similar formula for $\tilde{p}_k$. However, in a multi-matrix model, the moments $M_{k,l}$ are in general not functions of $k + l$.

Recursion relations

Since biorthogonal polynomials provide two bases of polynomials, there exist coefficients $Q_{n,m}$ and $\tilde{Q}_{n,m}$ such that

$$x p_n(x) = \sum_{m=0}^{n+1} Q_{n,m} \frac{\sqrt{h_n}}{\sqrt{h_m}} p_m(x) \ , \tag{5.229}$$

$$y \tilde{p}_n(y) = \sum_{m=0}^{n+1} \tilde{Q}_{n,m} \frac{\sqrt{h_n}}{\sqrt{h_m}} \tilde{p}_m(y) \ . \tag{5.230}$$

Then Q and $\tilde{Q}$ are two semi-infinite matrices, with at most one band above the diagonal, and

$$Q_{n,n+1} = \tilde{Q}_{n,n+1} = \gamma_{n+1} = \sqrt{\frac{h_{n+1}}{h_n}} \ . \tag{5.231}$$

In contrast to the matrix that appears in the recursion relation for orthogonal polynomials, the matrices Q and $\tilde{Q}$ are however not necessarily symmetric. Moreover, derivatives of biorthogonal polynomials are again linear combinations of biorthogonal polynomials,

$$p_n'(x) = \sum_{m=0}^{n-1} P_{n,m} \frac{\sqrt{h_n}}{\sqrt{h_m}} p_m(x) \ , \tag{5.232}$$

$$\tilde{p}_n'(y) = \sum_{m=0}^{n-1} \tilde{P}_{n,m} \frac{\sqrt{h_n}}{\sqrt{h_m}} \tilde{p}_m(y) \ . \tag{5.233}$$

Then P and $\tilde{P}$ are strictly lower triangular matrices, with

$$P_{n,n-1} = \tilde{P}_{n,n-1} = \frac{n}{\gamma_n} \ . \tag{5.234}$$

By definition of biorthogonal polynomials, the coefficients $P_{n,m}$ can be computed as scalar products, explicitly

$$P_{n,m} = \int_{\gamma\times\tilde{\gamma}} dxdy\ \omega(x,y) \frac{p'_n(x)}{\sqrt{h_n}} \frac{\tilde{p}_m(y)}{\sqrt{h_m}} \ . \tag{5.235}$$

Integrating by parts in the variable x leads to the expression of P in terms of Q and $\tilde{Q}$, and we find

$$\boxed{\begin{aligned} P &= -:\partial_x \log\omega(x,y)\big|_{(x,y)=(Q,\tilde{Q}^T)}: \ , & (5.236)\\ \tilde{P}^T &= -:\partial_y \log\omega(x,y)\big|_{(x,y)=(Q,\tilde{Q}^T)}: \ , & (5.237)\end{aligned}}$$

where $:f(x,y):$ is the function of the non-commuting variables x and y where y is inserted from the left and x from the right, for example $:(x+y)^2: := x^2 + 2yx + y^2$. Moreover, it follows immediately from their definitions that the matrices $Q, \tilde{Q}, P, \tilde{P}$ obey the string equations

$$[Q,P] = \mathrm{Id} \qquad , \qquad [\tilde{Q},\tilde{P}] = \mathrm{Id} \ . \tag{5.238}$$

Together with the expressions of $P, \tilde{P}$ as functions of $Q, \tilde{Q}$, this is in principle enough for determining these four infinite matrices.

Examples

- In the case of a two-matrix model (5.206), we find

$$P = V'(Q) - \tilde{Q}^T \ , \tag{5.239}$$

$$\tilde{P}^T = \tilde{V}'(\tilde{Q}^T) - Q \ . \tag{5.240}$$

  This not only reduces the string equations to

$$[\tilde{Q}^T, Q] = \mathrm{Id} \ , \tag{5.241}$$

  but also implies, since P is lower triangular,

$$\tilde{Q}_- = V'(Q^T)_- \ , \tag{5.242}$$

$$Q_- = \tilde{V}'(\tilde{Q}^T)_- \ . \tag{5.243}$$

  We know that Q and $\tilde{Q}$ have one non-vanishing band above the diagonal, and these equations now tell us what happens below the diagonal. In particular, if V is a polynomial of degree $d+1$, then $\tilde{Q}$ has at most d bands below the diagonal Therefore, if V and $\tilde{V}$ are polynomials, then all the matrices $Q, \tilde{Q}, P, \tilde{P}$ are band matrices – matrices with finitely many non-vanishing bands.

- In the rational case, which we define by the condition that $\partial_x \log\omega(x,y)$ and $\partial_y \log\omega(x,y)$ are rational fractions of their arguments, the string equations and the expressions for $P, \tilde{P}$ in terms of $Q, \tilde{Q}$ amount to recursion relations for the elements of these matrices, where the number of relations and the number of terms in each relation are N-independent. The space of solutions is finite-dimensional, and its dimension matches the N-dependent dimension of the homology space of contours on which the partition function converges. The matrices $Q, \tilde{Q}, P, \tilde{P}$ are not band matrices, but they can be written as algebraic combinations of band matrices.

- Let us consider the case

$$\omega(x,y) = e^{-f(xy)} \ . \tag{5.244}$$

This is in particular interesting if we integrate over the complex plane $\{(x,y) \in \mathbb{C}^2 | y = \bar{x}\}$ instead of a factorized contour $\gamma \times \tilde{\gamma}$, in which case we have a complex matrix model where $\omega(z,\bar{z}) = e^{-f(|z|^2)}$ depends only on the modulus. Now, using the symmetry $\omega(x,y) = \omega(y,x)$, we have $p_n = \tilde{p}_n$, $Q = \tilde{Q}$, and

$$P = \tilde{P} = - : \partial_x \log \omega(x,y)\big|_{(x,y)=(Q,Q^T)} := f'(QQ^T)Q^T \ . \tag{5.245}$$

For example, in the case $f' = 1$ which corresponds to the Gaussian complex matrix model, we have $P = -Q^T$ so that Q vanishes except on the band just above the diagonal, which implies

$$xp_n(x) = p_{n+1}(x) \qquad \text{and thus} \qquad p_n(x) = x^n \ . \tag{5.246}$$

We also have $h_n = \int_{\mathbb{C}} dx d\bar{x}\ x^n \bar{x}^n e^{-x\bar{x}} = \pi n!$.

- Let us finally consider a matrix chain (5.207), where the potentials V_i are polynomials of degrees $\deg V_i' = d_i$ [58, 59]. It can be proved that $Q, \tilde{Q}, P, \tilde{P}$ are band matrices, with Q having one band above and $d_2 d_3 \dots d_L$ bands below the diagonal, while $\tilde{Q}$ has one band above and $d_1 d_2 \dots d_{L-1}$ bands below the diagonal. In order to determine these matrices, it is useful to define the partial integrals

$$\psi_n^{(k)}(x_k) = \frac{1}{\sqrt{h_n}} \int dx_1 dx_2 \dots dx_{k-1}\, p_n(x_1) \prod_{i=1}^{k} e^{-V_i(x_i)} \prod_{i=1}^{k-1} e^{c_i x_i x_{i+1}} \ , \tag{5.247}$$

with in particular

$$\psi_n^{(1)}(x) = \frac{p_n(x)}{\sqrt{h_n}} e^{-V_1(x)} \ . \tag{5.248}$$

Let us consider the two families of infinite matrices $Q^{(k)}$ and $P^{(k)}$ such that

$$x\psi_n^{(k)}(x) = \sum_m Q_{n,m}^{(k)} \psi_m^{(k)}(x) \ , \tag{5.249}$$

$$\frac{d}{dx}\psi_n^{(k)}(x) = \sum_m P_{n,m}^{(k)} \psi_m^{(k)}(x) \ . \tag{5.250}$$

Of course, these matrices obey the string equations

$$[Q^{(k)}, P^{(k)}] = \mathrm{Id} \ . \tag{5.251}$$

Moreover, together with the analogously defined $\tilde{Q}^{(k)}$ and $\tilde{P}^{(k)}$, these matrices obey the relations

$$Q^{(k)} = \tilde{Q}^{(k)T} \qquad , \qquad P^{(k)} + \tilde{P}^{(k)T} = V_k'(Q^{(k)}) \ , \tag{5.252}$$

and

$$P^{(k)} = c_k \tilde{Q}^{(k+1)T} \qquad , \qquad \tilde{P}^{(k)} = c_{k-1} Q^{(k-1)T} \ . \tag{5.253}$$

These equations in principle determine $Q^{(k)}$, $\tilde{Q}^{(k)}$, $P^{(k)}$, and $\tilde{P}^{(k)}$, which are band matrices. From the matrices $Q^{(1)}$ are $\tilde{Q}^{(1)}$, we can then compute the biorthogonal polynomials.

5.4.3 Differential system

From the matrices Q and P that appear in the recursion relations and in the derivatives of the biorthogonal polynomials, we will now deduce the matrix differential equation that these biorthogonal polynomials obey. This equation is the basic object of the associated integrable system.

We assume that Q is a band matrix, with one band above and d bands below the diagonal. (This is the case in the matrix chain with polynomial potentials.) Then the recursion relations imply folding relations for $\psi_n = \psi_n^{(1)}$ (5.248),

$$\psi_n(x) = \sum_{m=N-d}^{N} (F^{(N)}(x))_{n,m} \psi_n(x) \ , \tag{5.254}$$

where the folding matrices $F^{(N)}(x)$ depends polynomially on x. More precisely, we have

$$F^{(N)}(x) = (L(x) - R(x))A^N \ , \tag{5.255}$$

where $L(x)$ is the upper triangular left inverse of $(x - Q)$, $R(x)$ is the lower triangular right inverse of $(x - Q)$, and the Christoffel–Darboux matrix $A^N = [\Pi_N, Q]$ is nonzero only in a square block of size $d + 1$,

$$A^N_{i,j} \neq 0 \qquad \Rightarrow \qquad \begin{cases} N - d \leq i \leq N \ , \\ N - 1 \leq j \leq N - 1 + d \ . \end{cases} \tag{5.256}$$

Then the vector $\vec{\psi}_N = (\psi_{N-d}, \dots, \psi_N)^T$ satisfies a differential system of order $d + 1$,

$$\vec{\psi}'_N(x) = D_N(x)\vec{\psi}_N(x) \ , \tag{5.257}$$

where the matrix $D_N(x)$ is the folding of the operator P,

$$D_N(x) = (\Pi_{N+1} - \Pi_{N-\tilde{d}+1})PF^{(N)}(x) \ . \tag{5.258}$$

The coefficients of $D_N(x)$ are polynomial functions of x, whose degrees are bounded by the number of bands of P. Therefore our differential system is isomonodromic.

Chapter 6

Angular integrals

In this Chapter, we consider matrix integrals that are not invariant under conjugation, and can therefore not be reduced to integrals over eigenvalues. Let us consider the integral over a Gaussian ensemble

$$\mathcal{Z} = \int_{E_N^\beta} dM \; e^{-\operatorname{Tr} MY} f_{\text{invariant}}(M) \; , \tag{6.1}$$

where Y is a constant matrix, and the function $f_{\text{invariant}}$ is invariant under conjugation. Using the behaviour (1.19) of the flat measure dM under the angular-radial decomposition $M = U\Lambda U^{-1}$, we find

$$\mathcal{Z} = \int_{\mathbb{R}^N} d\Lambda \; |\Delta(\Lambda)|^\beta \, f_{\text{invariant}}(\Lambda) \mathcal{Z}(\Lambda, Y) \; , \tag{6.2}$$

where we introduce the **angular integral** over the corresponding circular ensemble

$$\boxed{\mathcal{Z}(X,Y) = \int_{U_N^\beta} dU \, e^{-\operatorname{Tr} UXU^{-1}Y}} \; . \tag{6.3}$$

Since the Haar measure dU is invariant under the left and right actions of the group U_N^β on itself, we can always assume that X and Y are diagonal matrices, in which case our angular integral is called an **Itzykson–Zuber integral** [60]. On the other hand, it is also interesting to consider angular integrals where X and Y belong to the Lie algebra of U_N^β, which are called **Harish-Chandra integrals** [61]. These two types of integrals coincide in the case $\beta = 2$ of the ensemble U_N, because matrices in $\operatorname{Lie} U_N = iH_N$ are always diagonal up to conjugations by elements of U_N.

In angular integrals, integrands are not invariant under the left and right actions of the full group U_N^β on itself: there is only a residual invariance under the action of a subgroup. This residual invariance leads to constraints on the **moments**

$$\left\langle \prod_{k=1}^m U_{i_k,j_k} \prod_{l=1}^n U^{-1}_{j'_l,i'_l} \right\rangle = \frac{1}{\mathcal{Z}(X,Y)} \int_{U_N^\beta} dU \, e^{-\operatorname{Tr} UXU^{-1}Y} \prod_{k=1}^m U_{i_k,j_k} \prod_{l=1}^n U^{-1}_{j'_l,i'_l} \; . \tag{6.4}$$

In the U_N case, the residual invariance includes the left and right actions of the diagonal subgroup $(U_1)^N$, which act on matrix elements as $U_{i,j} \to \varphi_i U_{i,j} \psi_j$ with $|\varphi_i| = |\psi_j| = 1$. Then the moments must vanish unless $\{j_k\} = \{j'_l\}$ and $\{i_k\} = \{i'_l\}$, which implies in particular $m = n$. For example,

$$\langle U_{i,j} \rangle = 0 \; . \tag{6.5}$$

Non-vanishing moments are parametrized by pairs of permutations $\pi, \rho \in \mathfrak{S}_n$ such that $i_k = i'_{\pi(k)}$ and $j_l = j'_{\rho(l)}$.

This characterization of non-vanishing moments can be generalized to $\beta = 1, 4$. Instead of permutations, we would then encounter the elements of the Weyl group of U_N^β.

6.1 Harish-Chandra integrals

6.1.1 Harish-Chandra formula

Let us first give a more general formulation of Harish-Chandra integrals, which is valid not only for our circular ensembles U_N^β, but for arbitrary compact Lie groups G:

$$\boxed{\mathcal{Z}(X,Y) = \int_G dU\ e^{-\langle \mathrm{Ad}_U(X), Y\rangle}}\ , \tag{6.6}$$

where X and Y are elements of a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g} = \mathrm{Lie}\, G$, on which G acts by the adjoint group action, and $\langle X, Y\rangle$ is the Killing form. We assume that G is the maximal compact group in $\exp \mathfrak{g}$. The value of the integral is then given by the **Harish-Chandra formula** [61],

$$\boxed{\mathcal{Z}(X,Y) = \frac{C(G)}{\Delta(X)\Delta(Y)} \sum_{\sigma\in W} (-1)^\sigma e^{-\langle \sigma(X), Y\rangle}}\ , \tag{6.7}$$

where we introduce the following notations:

- $W = \mathrm{Weyl}(G)$ is the Weyl group of our matrix ensemble. The action of an element $\sigma \in W$ on $X \in \mathrm{Lie}(G)$ is denoted $\sigma(X)$. The signature of σ is defined as 1 or -1 depending on the parity of the length of σ, and denoted $(-1)^\sigma$.
- The **generalized Vandermonde determinant** is

  $$\Delta(X) = \prod_{\alpha>0} \alpha(X)\ , \tag{6.8}$$

  where the product runs over the positive roots of the Lie algebra.
- The constant $C(G)$ is

  $$C(G) = \prod_{j=1}^{\dim \mathfrak{h}} m_j! \prod_{\alpha>0} \frac{\langle \alpha, \alpha\rangle}{2}\ , \tag{6.9}$$

  where the m_j's are the exponents of the Lie algebra $\mathfrak{g}$. (The integers such that the eigenvalues of a Coxeter element of W when acting on $\mathfrak{h}$ are $e^{2\pi i \frac{m_j}{h}}$, where h is the Coxeter number.)

The Harish-Chandra formula can be proved with the help of a Duistermaat–Heckman localization argument [62]. This argument is applicable to Harish-Chandra integrals thanks to the invariance of $S(U) = -\langle \mathrm{Ad}_U(X), Y\rangle$ under the maximal torus action (in the U_N case, the left and right actions of $(U_1)^N$). This implies that the saddle point approximation turns out to be exact,

$$\int dU e^{S(U)} = \sum_{\sigma\in\{\text{saddle points}\}} \frac{e^{S(\sigma)}}{\det S''(\sigma)}\ . \tag{6.10}$$

The saddle points U of the Harish-Chandra integral obey $[\mathrm{Ad}_U(X), Y] = 0$, and are therefore parametrized by the elements of the Weyl group W, with $\mathrm{Ad}_{U_\sigma}(X) = \sigma(X)$. The eigenvalues of the Hessian matrix $S''(\sigma)$ are then $\alpha(\sigma(X))\alpha(Y)$ for α a positive root.

Let us now consider moments of the type $\langle f(\mathrm{Ad}_U(X), Y)\rangle$, where the function $f(X,Y)$ is assumed to obey $f(\mathrm{Ad}_U(X), \mathrm{Ad}_U(Y)) = f(X,Y)$. Such moments are given by the **generalized Harish-Chandra formula**,

$$\int_G dU\ e^{-\langle \mathrm{Ad}_U(X), Y\rangle} f(\mathrm{Ad}_U(X), Y) = \frac{\pi^{-\dim \mathfrak{n}_+(G)} \prod_{j=1}^{\dim \mathfrak{h}} m_j!}{\Delta(X)\Delta(Y)} \times \sum_{\sigma\in W} (-1)^\sigma e^{-\langle \sigma(X), Y\rangle} \int_{\mathfrak{n}_+(G)} dT\ e^{-\langle T, T^\dagger\rangle} f(\sigma(X)+T, Y+T^\dagger)\ , \quad (6.11)$$

where the complex nilpotent algebra $\mathfrak{n}_+(G) = \mathrm{Span}_{\mathbb{C}}\{E_\alpha\}_{\alpha>0}$ is generated by the root vectors E_α (elements of a Chevalley basis), and $(\lambda E_\alpha)^\dagger = \bar\lambda E_{-\alpha}$ for $\lambda\in\mathbb{C}$. The generalized Harish-Chandra formula converts an integral over G with its Haar measure, into a Gaussian integral over a linear space $\mathfrak{n}_+(G)$ with a standard Lebesgue measure. If f is polynomial, the Gaussian integral can then be evaluated using Wick's theorem. This is further facilitated by the fact that T is nilpotent.

6.1.2 Case of U_N

In the case $G = U_N$,

- X and Y are diagonal anti-Hermitian matrices, which can be parametrized by their eigenvalues $X = (X_i) \in (i\mathbb{R})^N$,
- the adjoint action is $\mathrm{Ad}_U(X) = UXU^{-1}$,
- the Weyl group is the group of permutations $W = \mathfrak{S}_N$, and $\sigma(X) = (X_{\sigma(i)})$,
- the positive roots are the forms $X_i - X_j$ with $i > j$, and the generalized Vandermonde determinant coincides with the Vandermonde determinant,
- $\mathfrak{n}_+(G) = \mathfrak{n}_+$ is the set of strictly upper triangular complex matrices, and $T^\dagger$ is the Hermitian conjugate of T.

In the Harish-Chandra formula, the sum over the Weyl group amounts to a determinant,

$$\mathcal{Z}(X,Y) = \frac{C(U_N)}{\Delta(X)\Delta(Y)} \det_{i,j} e^{-X_i Y_j}\ , \quad (6.12)$$

where $C(U_N) = \prod_{j=1}^{N-1} j!$, which is the Barnes function, due to $m_j = j$ and $\langle\alpha,\alpha\rangle = 2$. We will now compute the moments, using the generalized Harish-Chandra formula.

Morozov's formula for quadratic moments

We start with the moments $\langle |U_{i,j}|^2\rangle$, which we write as residues of a generating function,

$$\left\langle |U_{i,j}|^2 \right\rangle = \frac{1}{(2\pi i)^2} \oint_{X_i} dx \oint_{Y_j} dy\ \left\langle \mathrm{Tr}\, U \frac{1}{x-X} U^{-1} \frac{1}{y-Y} \right\rangle\ . \quad (6.13)$$

Using the generalized Harish-Chandra formula, we find

$$\left\langle \mathrm{Tr}\, U \frac{1}{x-X} U^{-1} \frac{1}{y-Y} \right\rangle = \frac{\pi^{-\frac{N(N-1)}{2}}}{\det_{i,j} e^{-X_i Y_j}} \times \sum_{\sigma\in\mathfrak{S}_N} (-1)^\sigma e^{-\langle\sigma(X), Y\rangle} \int_{\mathfrak{n}_+} dT\ e^{-\langle T, T^\dagger\rangle}\, \mathrm{Tr}\, \frac{1}{x-\sigma(X)-T}\, \frac{1}{y-Y-T^\dagger}\ . \quad (6.14)$$

Actually, $\frac{1}{x-X-T}$ is polynomial in T. This is shown by applying the formula $\frac{1}{1-M} = \sum_{k=0}^{\infty} M^k$, which holds for any matrix M such that $1-M$ is invertible, to the strictly upper triangular matrix $M = T\frac{1}{x-X}$:

$$\frac{1}{x-X-T} = \frac{1}{x-X}\sum_{k=0}^{N-1}\left(T\frac{1}{x-X}\right)^k , \tag{6.15}$$

where the sum stops at $k = N-1$ due to $T^N = 0$. This allows us to compute the matrix elements

$$\left(\frac{1}{x-X-T}\right)_{i,j} = \frac{1}{x-X_i}\sum_{k=0}^{N-1}\sum_{i=j_0<j_1<\cdots<j_{k-1}<j_k=j}\prod_{l=1}^{k} T_{j_{l-1},j_l}\frac{1}{x-X_{j_l}} . \tag{6.16}$$

We can now use Wick's theorem for computing the integral over T, using the two-point function

$$\pi^{-\frac{N(N-1)}{2}}\int_{\mathfrak{n}_+} dT\; e^{-\langle T,T^\dagger\rangle} T_{i,j}T^\dagger_{k,l} = \delta_{i,l}\delta_{j,k} . \tag{6.17}$$

Using Wick's theorem is particularly simple because we are integrating objects $T^k = \prod_{l=1}^{k} T_{j_{l-1},j_l}$ with ordered indices $j_0 < j_1 < \cdots < j_k$. The integral of $T^k(T^\dagger)^k$ therefore involves at most one pairing, and we find

$$\begin{aligned}
&\pi^{-\frac{N(N-1)}{2}}\int_{\mathfrak{n}_+} dT\; e^{-\langle T,T^\dagger\rangle}\,\mathrm{Tr}\,\frac{1}{x-\sigma(X)-T}\frac{1}{y-Y-T^\dagger}\\
&= \sum_{k=0}^{N-1}\sum_{j_0<j_1<\cdots<j_{k-1}<j_k}\prod_{l=0}^{k}\frac{1}{(x-X_{\sigma(j_l)})(y-Y_{j_l})}\\
&= -1+\prod_{j=1}^{N}\left(1+\frac{1}{(x-X_{\sigma(j)})(y-Y_j)}\right) .
\end{aligned} \tag{6.18}$$

We therefore have

$$1+\left\langle \mathrm{Tr}\,U\frac{1}{x-X}U^{-1}\frac{1}{y-Y}\right\rangle = \frac{\det_{i,j} e^{-X_iY_j}\left(1+\frac{1}{(x-X_i)(y-Y_j)}\right)}{\det_{i,j} e^{-X_iY_j}} , \tag{6.19}$$

$$= \frac{\det\left(E+\frac{1}{x-X}E\frac{1}{y-Y}\right)}{\det E} , \tag{6.20}$$

$$= \det\left(1+E^{-1}\frac{1}{x-X}E\frac{1}{y-Y}\right) , \tag{6.21}$$

where we introduced the matrix

$$\boxed{E_{i,j} = e^{-X_iY_j}} . \tag{6.22}$$

We can now extract our moments as residues (6.13), and we find **Morozov's formula** [63]

$$\boxed{\left\langle |U_{i,j}|^2\right\rangle = \frac{\mathrm{Minor}_{i,j}\left(E+\frac{1}{X_i-X}E\frac{1}{Y_j-Y}\right)}{\det E}} . \tag{6.23}$$

Higher moments

Similarly, we can compute moments of $2n$ matrix elements for any integer n, following [64]. (See [65] for the generalization to arbitrary Lie groups.) We parametrize the nonvanishing moments $\left\langle U_{i_1,j_{\pi(1)}} \dots U_{i_n,j_{\pi(n)}} U^{-1}_{j_{\rho(1)},i_1} \dots U^{-1}_{j_{\rho(n)},i_n} \right\rangle$ with two sets of indices $\{i_1, \dots i_n\} \subset \{1, \dots N\}$ and $\{j_1, \cdots j_n\} \subset \{1, \dots N\}$, and two permutations $\pi, \rho \in \mathfrak{S}_n$. We will however trade the indices $\{i_1, \dots i_n\}$ and $\{j_1, \cdots j_n\}$ for complex variables $\vec{x} = (x_1, \dots x_n)$, $\vec{y} = (y_1, \dots y_n)$, using the transformation

$$\begin{aligned} &U_{i_1,j_{\pi(1)}} \dots U_{i_n,j_{\pi(n)}} U^{-1}_{j_{\rho(1)},i_1} \dots U^{-1}_{j_{\rho(n)},i_n} \\ &\quad = \frac{1}{(2\pi i)^{2n}} \oint_{X_{i_1}} dx_1 \dots \oint_{X_{i_n}} dx_n \oint_{Y_{i_1}} dy_1 \dots \oint_{Y_{i_n}} dy_n \; H(\vec{x}, \vec{y}; UXU^{-1}, Y)_{\pi,\rho} \;, \end{aligned} \tag{6.24}$$

where we introduced the generating function

$$\boxed{H(\vec{x}, \vec{y}; X, Y)_{\pi,\rho} = \prod_{c=\text{cycle of } \pi\circ\rho^{-1}} \left(\delta_{\text{length}(c),1} + \text{Tr} \prod_{i\in c} \frac{1}{x_{\rho(i)} - X} \frac{1}{y_i - Y} \right)} \,. \tag{6.25}$$

Here the term $\delta_{\text{length}(c),1}$ does not contribute to the moments. Including this term, which is similar to the term 1 in eq. (6.19), will however lead to simpler formulas. For example, in the case $n = 2$,

$$H(\vec{x}, \vec{y}; X, Y) = \begin{bmatrix} \left(1+\text{Tr}\, \frac{1}{x_1-X} \frac{1}{y_1-Y}\right)\left(1+\text{Tr}\, \frac{1}{x_2-X} \frac{1}{y_2-Y}\right) & \text{Tr}\, \frac{1}{x_1-X} \frac{1}{y_1-Y} \frac{1}{x_2-X} \frac{1}{y_2-Y} \\ \text{Tr}\, \frac{1}{x_1-X} \frac{1}{y_2-Y} \frac{1}{x_2-X} \frac{1}{y_1-Y} & \left(1+\text{Tr}\, \frac{1}{x_1-X} \frac{1}{y_2-Y}\right)\left(1+\text{Tr}\, \frac{1}{x_2-X} \frac{1}{y_1-Y}\right) \end{bmatrix} . \tag{6.26}$$

Using Wick's theorem, it is possible to compute

$$\boxed{\left\langle H(\vec{x}, \vec{y}; UXU^{-1}, Y) \right\rangle = \frac{1}{\det E} \sum_{\sigma\in\mathfrak{S}_N} (-1)^\sigma \prod_{i=1}^N M(\vec{x}, \vec{y}; X_{\sigma(i)}, Y_i)} \,, \tag{6.27}$$

where we introduced the $n! \times n!$ matrix $M(\vec{x}, \vec{y}; x, y)$, which depends on auxiliary complex variables x and y, and whose elements are

$$\boxed{M(\vec{x}, \vec{y}; x, y)_{\pi,\rho} = e^{-xy} \prod_{i=1}^n \left(\delta_{\pi(i),\rho(i)} + \frac{1}{(x - x_{\rho(i)})(y - y_i)} \right)} \,. \tag{6.28}$$

For example, in the case $n = 2$,

$$M(\vec{x}, \vec{y}; x, y) = e^{-xy} \begin{bmatrix} \left(1+\frac{1}{(x-x_1)(y-y_1)}\right)\left(1+\frac{1}{(x-x_2)(y-y_2)}\right) & \frac{1}{(x-x_1)(x-x_2)(y-y_1)(y-y_2)} \\ \frac{1}{(x-x_1)(x-x_2)(y-y_1)(y-y_2)} & \left(1+\frac{1}{(x-x_1)(y-y_2)}\right)\left(1+\frac{1}{(x-x_2)(y-y_1)}\right) \end{bmatrix} . \tag{6.29}$$

The matrix $M(\vec{x}, \vec{y}; x, y)$ enjoys beautiful properties:

$$\text{Symmetry:} \qquad M(\vec{x}, \vec{y}; x, y) = M(\vec{x}, \vec{y}; x, y)^T \;, \tag{6.30}$$

$$\text{Commutativity:} \qquad \left[M(\vec{x}, \vec{y}; x, y), M(\vec{x}, \vec{y}; x', y')\right] = 0 \;. \tag{6.31}$$

In particular, commutativity implies that matrix product in eq. (6.27) does not depend on the order. In the case $n = 1$, we recover Morozov's formula.

6.1.3 Application to matrix chains

We will now show how the expressions of Hermitian matrix chain partition functions as eigenvalue integrals that we used in Section 5.4 can be derived using the Harish-Chandra formula.

One-matrix model with an external field

Let us write the partition function of the Hermitian one-matrix model with an external field as

$$\mathcal{Z}(A) = \int_{H_N(\gamma)} dM\, e^{-\operatorname{Tr} V(M) + \operatorname{Tr} MA} . \tag{6.32}$$

Without loss of generality, we assume that the external field is diagonal, so that $A = \operatorname{diag}(a_1, \dots, a_N)$. Let us use the angular-radial decomposition $M = U\Lambda U^{-1}$ with $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_N)$, and perform the angular integral with the help of the Harish-Chandra formula (6.12). Neglecting A, V-independent prefactors, we find

$$\mathcal{Z}(A) \propto \frac{1}{N!\Delta(A)} \int_{\gamma^N} d\Lambda\, e^{-\operatorname{Tr} V(\Lambda)} \Delta(\Lambda) \det_{1\le i,j\le N} e^{\lambda_i a_j} . \tag{6.33}$$

Using the expression (5.2) of the Vandermonde determinant $\Delta(\Lambda)$, and the identity

$$\frac{1}{N!} \int d^N\lambda \left(\det_{1\le i,j\le N} f_i(\lambda_j) \right) \left(\det_{1\le i,j\le N} g_i(\lambda_j) \right) = \det_{1\le i,j\le N} \left(\int d\lambda\, f_i(\lambda) g_j(\lambda) \right) , \tag{6.34}$$

we deduce

$$\mathcal{Z}(A) \propto \frac{1}{\Delta(A)} \det_{1\le i,j\le N} \int_\gamma d\lambda\, p_{i-1}(\lambda) e^{-V(\lambda)+\lambda a_j} , \tag{6.35}$$

for any family of monic polynomials $p_i(\lambda) = \lambda^i + \cdots$. In particular, in the case $p_i(\lambda) = \lambda^i$, we can take the limit $A \to 0$ and recover the expression (5.8) of the partition function of the one-matrix model as a Hankel determinant. (The singularity of the $\frac{1}{\Delta(A)}$ prefactor is cancelled by the determinant, as can be seen using the Taylor expansion at $A = 0$.)

Two-matrix model

In the Hermitian two-matrix model with the potentials V_1 and V_2, let us use the angular-radial decompositions of the two matrices, $M_1 = U_1\Lambda U_1^{-1}$, $M_2 = U_2\tilde{\Lambda}U_2^{-1}$ with $\Lambda = \operatorname{diag}(\lambda_1, \dots \lambda_N)$, $\tilde{\Lambda} = \operatorname{diag}(\tilde{\lambda}_1, \dots \tilde{\lambda}_N)$. The Harish-Chandra formula allows us to perform the angular integrals,

$$\int_{(U_N)^2} dU_1 dU_2\, e^{\operatorname{Tr} U_1\Lambda U_1^{-1} U_2\tilde{\Lambda}U_2^{-1}} \propto \frac{1}{\Delta(\Lambda)\Delta(\tilde{\Lambda})} \det_{1\le i,j\le N} e^{\lambda_i\tilde{\lambda}_j} , \tag{6.36}$$

and the partition function can be written as the radial integral

$$\mathcal{Z} \propto \int d\Lambda d\tilde{\Lambda}\, e^{-\operatorname{Tr}\left(V_1(\Lambda)+V_2(\tilde{\Lambda})\right)} \Delta(\Lambda)\Delta(\tilde{\Lambda}) \det_{1\le i,j\le N} e^{\lambda_i\tilde{\lambda}_j} , \tag{6.37}$$

in agreement with eq. (5.205).

Matrix chain

In the Hermitian matrix chain of length L with no external field, a similar calculation leads to the representation of the partition function as a radial integral,

$$\mathcal{Z} \propto \int \prod_{i=1}^{L} d\Lambda_i \, e^{-\operatorname{Tr} V_i(\Lambda^{(i)})} \Delta(\Lambda^{(1)}) \Delta(\Lambda^{(L)}) \prod_{i=1}^{L-1} \left(\det_{1 \leq j,k \leq N} e^{\lambda_j^{(i)} \lambda_k^{(i+1)}} \right) . \tag{6.38}$$

Applying the formula (6.34) to the integrals over $\Lambda^{(2)}, \dots \Lambda^{(L-1)}$, we recover eq. (5.205).

6.2 Itzykson–Zuber integrals

Although there is no simple expression for Itzykson–Zuber integrals [60], there are a many known relations, which we will now review following [66]. These relations can be used for defining Itzykson–Zuber integrals for arbitrary complex values of β, beyond their original definition as matrix integrals for $\beta \in \{1, 2, 4\}$.

6.2.1 Calogero–Moser equation

We will show that the Itzykson–Zuber integral $\mathcal{Z}(X, Y)$ obeys second-order differential equations. We will not do a full proof, but begin with first-order equations for the quadratic moments

$$M_{i,j} = \mathcal{Z}(X, Y) \left\langle U_{i,j} U_{j,i}^{-1} \right\rangle = \int_{U_N^\beta} dU \; e^{-\operatorname{Tr} UXU^{-1}Y} U_{i,j} U_{j,i}^{-1} \; , \tag{6.39}$$

which obey the trivial equations

$$\sum_{i=1}^{N} M_{i,j} = \sum_{j=1}^{N} M_{i,j} = \mathcal{Z}(X, Y) \, , \qquad \frac{\partial \mathcal{Z}(X, Y)}{\partial X_i} = -\sum_{j=1}^{N} M_{i,j} Y_j \; . \tag{6.40}$$

We now accept that, as a consequence of the loop equations, the quadratic moments obey the following first-order differential equations,

$$\frac{\partial M_{i,j}}{\partial X_i} = -M_{i,j} Y_j - \frac{\beta}{2} \sum_{k \neq i} \frac{M_{i,j} - M_{k,j}}{X_i - X_k} \; , \tag{6.41}$$

which involve differential operators called Dunkl operators. We then compute

$$\sum_i \frac{\partial^2}{\partial X_i^2} \mathcal{Z}(X, Y) = -\sum_{i,j} \frac{\partial}{\partial X_i} M_{i,j} Y_j \; , \tag{6.42}$$

$$= \sum_{i,j} M_{i,j} Y_j^2 + \frac{\beta}{2} \sum_j \sum_{k \neq i} \frac{M_{i,j} Y_j - M_{k,j} Y_j}{X_i - X_k} \; , \tag{6.43}$$

which leads to

$$\boxed{\left\{ \sum_i \frac{\partial^2}{\partial X_i^2} + \frac{\beta}{2} \sum_{k \neq i} \frac{1}{X_i - X_k} \left(\frac{\partial}{\partial X_i} - \frac{\partial}{\partial X_k} \right) \right\} \mathcal{Z}(X, Y) = \sum_j Y_j^2 \mathcal{Z}(X, Y)} \, . \tag{6.44}$$

The interpretation is that $\mathcal{Z}(X,Y)$, as a function of X, is an eigenfunction of a **Calogero–Moser system** of N second-order differential operators, with the eigenvalues $\sum_j Y_j^2$. Actually, $\mathcal{Z}(X,Y)$ can be characterized as the eigenfunction with the additional properties

$$\mathcal{Z}(X,Y) = \mathcal{Z}(Y,X) \ , \tag{6.45}$$

$$\forall \sigma \in W \ , \quad \mathcal{Z}(\sigma(X),Y) = \mathcal{Z}(X,Y) \ , \tag{6.46}$$

where $\sigma(X)$ is the action of the Weyl group element σ on the maximal torus to which X and Y belong.

6.2.2 Jack polynomials

The Itzykson–Zuber integral $\mathcal{Z}(X,Y)$ is an eigenfunction of the Calogero–Moser Hamiltonian, which is a symmetric function of the components X_i. This implies that $\mathcal{Z}(X,Y)$ can be written as a linear combination of the Jack polynomials, as these polynomials form a basis of symmetric polynomials. We will now review the definition and properties of these polynomials, while referring to the textbook [67] for more details.

For any partition $\lambda = (\lambda_1, \dots, \lambda_N)$ of length N, we define the symmetrized monomial of N variables

$$m_\lambda(X_1, \dots, X_N) = \frac{1}{\prod_k n_k(\lambda)!} \sum_{\sigma \in \mathfrak{S}_N} \prod_{i=1}^N X_{\sigma(i)}^{\lambda_i} \ , \tag{6.47}$$

where we introduced

$$n_k(\lambda) = \#\{i, \ \lambda_i = k\} \, . \tag{6.48}$$

We also define a partial ordering on partitions of length N by

$$\lambda \leq \mu \qquad \Leftrightarrow \qquad \forall 1 \leq k \leq N \ , \quad \sum_{i=1}^k \lambda_i \leq \sum_{i=1}^k \mu_i \ . \tag{6.49}$$

And for any number α we define a scalar product of two symmetric functions of N variables by

$$\langle f | g \rangle_\alpha = \int \prod_{i=1}^N \frac{dX_i}{X_i} \, |\Delta(X)|^{\frac{1}{\alpha}} \left|\Delta(X^{-1})\right|^{\frac{1}{\alpha}} f(X) g(X^{-1}) \ . \tag{6.50}$$

Then the Jack polynomials $J_\lambda^{(\alpha)}$ are the homogeneous symmetric polynomials such that

$$J_\lambda^{(\alpha)}(X) = m_\lambda(X) + \sum_{\substack{\mu < \lambda \\ |\mu| = |\nu|}} c_{\lambda,\mu} m_\mu(X) \quad \text{and} \quad \left\langle J_\lambda^{(\alpha)} \middle| J_\mu^{(\alpha)} \right\rangle_\alpha \neq 0 \Rightarrow \mu = \nu \ , \tag{6.51}$$

where $|\lambda|$ is the weight of the partition λ. So two different Jack polynomials are orthogonal to each other, and the squared norm of a Jack polynomial is

$$\left\langle J_\lambda^{(\alpha)} \middle| J_\lambda^{(\alpha)} \right\rangle_\alpha = \prod_{j=1}^N \frac{\Gamma((N-j+1)\alpha^{-1} + \lambda_j)\Gamma((N-j)\alpha^{-1}+1)}{\Gamma((N-j+1)\alpha^{-1})\Gamma((N-j)\alpha^{-1}+\lambda_j+1)} \, . \tag{6.52}$$

Moreover, Jack polynomials provide a simple decomposition of the function

$$\frac{1}{\prod_{i=1}^N \prod_{j=1}^N (1 - X_i Y_j)^{\frac{1}{\alpha}}} = \sum_\lambda J_\lambda^{(\alpha)}(X) J_\lambda^{(\alpha)}(Y) \, . \tag{6.53}$$

The reason why Jack polynomials are relevant to Itzykson–Zuber integrals is that they are eigenfunctions of the Calogero–Moser system with the parameter $\beta = \frac{2}{\alpha}$. The Itzykson–Zuber integral is then expressed in terms of Jack polynomials as [68]

$$\boxed{\mathcal{Z}(X,Y) = \sum_{\lambda} (-1)^{|\lambda|} d_\lambda J_\lambda^{(\frac{2}{\beta})}(X) J_\lambda^{(\frac{2}{\beta})}(Y) \ , \quad d_\lambda = \prod_{j=1}^{N} \frac{\Gamma((N-j+1)\frac{\beta}{2})}{\Gamma((N-j+1)\frac{\beta}{2}+\lambda_j)}} \ . \tag{6.54}$$

In contrast to the Harish-Chandra formula, this formula involves an infinite sum, which moreover does not converge fast.

6.2.3 Lagrange multipliers and duality equation

Let us realize the elements $U \in U_N^\beta$ of our circular ensemble as size-N matrices subject to quadratic constraints, as we did in Section 1.2.1. This will allow us to rewrite the corresponding matrix integrals in terms of integrals on $\mathbb{R}^{\beta N^2}$ with the Lebesgue measure, at the expense of introducing Lagrange multipliers for imposing the constraints.

Let $U_{i,j}$ be the N^2 coefficients of U, and $v_j = (U_{i,j})_{i=1,\dots N}$ be the column vectors. The orthonormality constraint $UU^\dagger = \mathrm{Id}$ amounts to $v_i^\dagger v_j = \delta_{i,j}$, let $S_{j,i} = \bar{S}_{i,j}$ be the corresponding Lagrange multipliers. We then have

$$\mathcal{Z}(X,Y) = \int dv \ e^{-\sum_{i,j} U_{i,j} X_i \bar{U}_{i,j} Y_j} \prod_{i,j} \delta(\delta_{i,j} - v_i^\dagger v_j) \ , \tag{6.55}$$

$$= \int dv \ e^{-\sum_j Y_j v_j^\dagger (X v_j)} \int dS \ e^{i \sum_{i,j} S_{j,i} (\delta_{i,j} - v_i^\dagger v_j)} \ , \tag{6.56}$$

$$= \int dS \ e^{i \operatorname{Tr} S} \int dv \ e^{-\sum_j Y_j v_j^\dagger (X v_j)} e^{-i \sum_{i,j} S_{j,i} v_i^\dagger v_j} \ , \tag{6.57}$$

where the measures dS and dv are products of Lebesgue measures on the real components $U_{i,j}^{(\alpha)}$ and $S_{i,j}^{(\alpha)}$. Performing the Gaussian integral over v, and doing the rescaling $S \to Y^{\frac{1}{2}} S Y^{\frac{1}{2}}$, we obtain

$$\mathcal{Z}(X,Y) = \frac{\pi^{\frac{\beta}{2}N^2}}{\det Y^{\frac{\beta}{2}-1}} \int dS \ \frac{e^{i \operatorname{Tr} SY}}{\prod_{j=1}^N \det(iS + X_j)^{\frac{\beta}{2}}} \ . \tag{6.58}$$

In this formulation of the matrix integral, the Haar measure dU has been replaced with the flat measure dS. The resulting integral is itself of the type of eq. (6.1), so that diagonalizing S leads to an integral over its real eigenvalues, and an angular integral which is again an Itzykson–Zuber integral. Therefore, $\mathcal{Z}(X,Y)$ obeys the duality equation

$$\boxed{\mathcal{Z}(X,Y) = \frac{\pi^{\frac{\beta}{2}N^2}}{\det Y^{\frac{\beta}{2}-1}} \int_{\mathbb{R}^N} d\Lambda \ \frac{|\Delta(\Lambda)|^\beta}{\prod_{j=1}^N \det(i\Lambda + X_j)^{\frac{\beta}{2}}} \ \mathcal{Z}(i\Lambda, Y)} \ . \tag{6.59}$$

So, as a function of X, the combination $|\Delta(X)|^{\frac{\beta}{2}} \mathcal{Z}(X,Y)$ is covariant under convolution with the kernel $\left| \frac{\Delta(i\Lambda)\Delta(X)}{\prod_{j,k}(i\Lambda_j + X_k)} \right|^{\frac{\beta}{2}}$, which is a power of a Cauchy determinant. This duality can be checked to be compatible with the expression (6.54) of $\mathcal{Z}(X,Y)$ in terms of Jack polynomials, using their property (6.53).

6.2.4 Lagrange multipliers and recursion equation

Let us come back to the expression (6.55) for $\mathcal{Z}_N(X,Y) = \mathcal{Z}(X,Y)$ as an integral over a basis $v = (v_j)$, and deduce a recursion equation for the dependence on the matrix size N. To do this, we will integrate over the last vector v_N before introducing Lagrange multipliers. That integral is easy to do, because the orthonormality constraints determine the value of v_N as a function of $v_1, \dots v_{N-1}$, up to an element of U_1^β. Using a shift of Y to further eliminate all dependence on v_N from the integrand, we compute

$$\mathcal{Z}_N(X,Y) = e^{-Y_N \operatorname{Tr} X} \mathcal{Z}_N(X, Y - Y_N \mathrm{Id}) \ , \tag{6.60}$$

$$= \mathrm{Vol}(U_1^\beta) e^{-Y_N \operatorname{Tr} X} \int \prod_{j=1}^{N-1} \left(dv_j \ e^{-(Y_j - Y_N) v_j^\dagger (X v_j)} \right) \prod_{i,j=1}^{N-1} \delta(\delta_{i,j} - v_i^\dagger v_j) \ , \tag{6.61}$$

where

$$\mathrm{Vol}(U_1^\beta) = \frac{2\pi^{\frac{\beta}{2}}}{\Gamma\left(\frac{\beta}{2}\right)} \ . \tag{6.62}$$

Now we only need Lagrange multipliers for $(N-1)^2$ constraints. The angular part of the integral over the Lagrange multipliers is therefore an Itzykson–Zuber integral over matrices of size $N-1$,

$$\mathcal{Z}_N(X,Y) \propto \frac{e^{-Y_N \operatorname{Tr} X}}{\prod_{i=1}^{N-1} (Y_i - Y_N)^{\frac{\beta}{2}-1}} \int dS \ \frac{e^{i \operatorname{Tr} S(Y - Y_N \mathrm{Id})}}{\prod_{j=1}^{N} \det(iS + X_j)^{\frac{\beta}{2}}} \ , \tag{6.63}$$

and we end up with the recursion equation

$$\boxed{\mathcal{Z}_N(X,Y) \propto \frac{e^{-Y_N \operatorname{Tr} X}}{\prod_{i=1}^{N-1} (Y_i - Y_N)^{\frac{\beta}{2}-1}} \int_{\mathbb{R}^{N-1}} d\Lambda \ \frac{|\Delta(\Lambda)|^\beta e^{i Y_N \operatorname{Tr} \Lambda}}{\prod_{j=1}^{N} \det(i\Lambda + X_j)^{\frac{\beta}{2}}} \mathcal{Z}_{N-1}(i\Lambda, Y)} \ , \tag{6.64}$$

where $\mathcal{Z}_{N-1}(i\Lambda, Y)$ depends on the truncated diagonal matrix $(Y_1, \dots Y_{N-1})$, and we omitted X, Y-independent prefactors.

Analytic properties of Itzykson–Zuber integrals

Let us assume that β is an even integer. Then the integral over Λ_j is a sum of residues at $\Lambda_j = iX_i$. The integral over Λ is therefore a sum of $(N-1)$-dimensional residues at $(\Lambda_j) = (iX_{\sigma(j)})$, parametrized by maps $\sigma : \{1, \dots N-1\} \to \{1, \dots N\}$. Due to the factor $|\Delta(\Lambda)|^\beta$, only injective maps give nonzero contributions. This can be used not only for computing Itzykson–Zuber integrals, but also for deducing their analytic properties. Namely, for any value of N, there exists a rational function $\hat{\mathcal{Z}}_N(X,Y)$ such that

$$\boxed{\mathcal{Z}_N(X,Y) = \sum_{\sigma \in W} \frac{\prod_{i=1}^N e^{-\sum_i X_i Y_{\sigma(i)}}}{|\Delta(X)\Delta(\sigma(Y))|^{\frac{\beta}{2}}} \hat{\mathcal{Z}}_N(X, \sigma(Y))} \ . \tag{6.65}$$

In the case $\beta = 2$, that rational function is a constant, and we recover the Harish-Chandra formula. More generally, that rational function is a symmetric polynomial of degree β in the variables

$$\tau_{i,j} = -\frac{2}{(X_i - X_j)(Y_i - Y_j)} \ . \tag{6.66}$$

These $\frac{1}{2}N(N-1)$ variables are redundant for $N \geq 6$, since they are combinations of the $2N$ variables X_i, Y_i.

Actually, even when β is not an even integer, the singularities of $\mathcal{Z}_N(X,Y)$ are described by the same formula, but the factor $\hat{\mathcal{Z}}_N(X,Y)$ is an analytic function of the variables $\tau_{i,j}$ instead of a polynomial.

Special cases $N = 1, 2, 3$

As an analytic function of $\tau_{i,j}$, the factor $\hat{\mathcal{Z}}_N(X,Y)$ can be written in terms of the analytic function

$$y_\alpha(x) = \sum_{k=0}^{\infty} \frac{\Gamma(\alpha+k+1)}{k!\Gamma(\alpha-k+1)} \left(\frac{x}{2}\right)^k = \sqrt{\frac{2}{\pi x}} e^{\frac{1}{x}} K_{\alpha+\frac{1}{2}}\left(\frac{1}{x}\right) \ , \tag{6.67}$$

which is the Bessel polynomial of degree α if $\alpha \in \mathbb{N}$, and is in general related to the modified Bessel function of the second kind $K_{\alpha+\frac{1}{2}}$. Starting with the trivial case $N = 1$ and using the recursion equation, we find

$$\hat{\mathcal{Z}}_1(X,Y) = 1 \ , \tag{6.68}$$

$$\hat{\mathcal{Z}}_2(X,Y) = y_{\frac{\beta}{2}-1}(\tau_{1,2}) \ , \tag{6.69}$$

$$\hat{\mathcal{Z}}_3(X,Y) = \sum_{k=0}^{\infty} \frac{\Gamma(\frac{\beta}{2}-k)}{2^{6k}k!\Gamma(\frac{\beta}{2}+k)} \prod_{1\leq i<j\leq 3} y^{(k)}_{\frac{\beta}{2}-1}(\tau_{i,j}) \ , \tag{6.70}$$

where $y^{(k)}_{\frac{\beta}{2}-1}$ is the k-th derivative of $y_{\frac{\beta}{2}-1}$, which vanishes for $k \geq \frac{\beta}{2}$ if β is an even integer.

Bibliography

[1] M. L. Mehta, *Random Matrices*, vol. 142 of *Pure and Applied Mathematics*. Academic Press, 3rd ed., 2004.

[2] P. Bleher and A. Its, eds., *Random Matrix Models and their Applications*. Cambridge Univ. Press, 2001.

[3] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics. Cambridge Univ. Press, 2009.

[4] P. J. Forrester, *Log-Gases and Random Matrices*, vol. 34 of *London Mathematical Society Monographs*. Princeton Univ. Press, 2010.

[5] G. Akemann, J. Baik, and P. Di Francesco, eds., *The Oxford Handbook of Random Matrix Theory*. Oxford Handbooks in Mathematics. Oxford Univ. Press, 2011.

[6] J. Harnad, ed., *Random Matrices, Random Processes and Integrable Systems*. CRM Series in Mathematical Physics. Springer New York, 2011.

[7] V. A. Kazakov, M. Staudacher, and T. Wynter, "Character expansion methods for matrix models of dually weighted graphs," *Commun. Math. Phys.* **177** (1996) 451–468, arXiv:hep-th/9502132 [hep-th].

[8] K. Efetov, *Supersymmetry in Disorder and Chaos*. Cambridge Univ. Press, 1996.

[9] T. Tao and V. Vu, "Random Matrices: Universality of Local Eigenvalue Statistics up to the Edge," *Commun. Math. Phys.* **298** (2010) 549–572, arXiv:0908.1982 [math.PR].

[10] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population," *Biometrika* **20A** (1928) 32–52.

[11] V. A. Marchenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mat. Sb.* **72** (1967) 507–536.

[12] M. V. Berry and M. Tabor, "Level Clustering in the Regular Spectrum," *Proc. Roy. Soc.* **A356** (1977) 375–394.

[13] O. Bohigas, M. J. Giannoni, and C. Schmit, "Characterization of Chaotic Quantum Spectra and Universality of Level Fluctuation Laws," *Phys. Rev. Lett.* **52** (1984) 1–4.

[14] G. 't Hooft, "A Planar Diagram Theory for Strong Interactions," *Nucl. Phys.* **B72** (1974) 461–473.

[15] E. Brézin, C. Itzykson, G. Parisi, and J.-B. Zuber, "Planar Diagrams," *Commun. Math. Phys.* **59** (1978) 35–51.

[16] R. Dijkgraaf and C. Vafa, "Matrix models, topological strings, and supersymmetric gauge theories," *Nucl. Phys.* **B644** (2002) 3–20, arXiv:hep-th/0206255.

[17] M. Kontsevich, "Intersection theory on the moduli space of curves and the matrix Airy function," *Commun. Math. Phys.* **147** (1992) 1–23.

[18] V. Bouchard, A. Klemm, M. Mariño, and S. Pasquetti, "Remodeling the B-model," *Commun. Math. Phys.* **287** (2009) 117–178, arXiv:0709.1453 [hep-th].

[19] B. Eynard and N. Orantin, "Computation of open Gromov–Witten invariants for toric Calabi–Yau 3-folds by topological recursion, a proof of the BKMP conjecture," *Commun. Math. Phys.* **337** (2015) 483–567, arXiv:1205.1103 [math-ph].

[20] R. C. Penner, "Perturbative series and the moduli space of Riemann surfaces," *J. Differential Geom.* **27** (1988) 35–53.

[21] V. S. Dotsenko and V. A. Fateev, "Conformal algebra and multipoint correlation functions in 2D statistical models," *Nucl. Phys.* **B240** (1984) 312–348.

[22] V. S. Dotsenko and V. A. Fateev, "Four-point correlation functions and the operator algebra in 2D conformal invariant theories with central charge $c \leq 1$," *Nucl. Phys.* **B251** (1985) 691–734.

[23] B. Eynard, "A Matrix model for plane partitions and (T)ASEP," *J.Stat.Mech.* **0910** (2009) P10011, arXiv:0905.0535 [math-ph].

[24] M. Mineev-Weinstein, P. B. Wiegmann, and A. Zabrodin, "Integrable Structure of Interface Dynamics," *Phys. Rev. Lett.* **84** (2000) 5106–5109, arXiv:nlin/0001007 [nlin-si].

[25] I. Kostov, I. Krichever, M. Mineev-Weinstein, P. Wiegmann, and A. Zabrodin, "τ-function for analytic curves," in *Random Matrix Models and their Applications*, vol. 40 of *Math. Sci. Res. Inst. Publ.*, pp. 285–299. Cambridge Univ. Press, 2001. arXiv:hep-th/0005259 [hep-th].

[26] H. L. Montgomery, "The pair correlation of zeros of the zeta function," *Analytic Number Theory, Proc. Sympos. Pure Math.* **XXIV** (1973) 181–193.

[27] A. M. Odlyzko, "On the distribution of spacings between zeros of the zeta function," *Math. Comp.* **48** (1987) 273–308.

[28] H. Orland and A. Zee, "RNA Folding and Large N Matrix Theory," *Nucl. Phys.* **B620** (2002) 456–476, arXiv:cond-mat/0106359 [cond-mat].

[29] A. Brini, B. Eynard, and M. Marino, "Torus knots and mirror symmetry," *Annales Henri Poincare* **13** (2012) 1873–1910, arXiv:1105.2012 [hep-th].

[30] M. Caselle and U. Magnea, "Random matrix theory and symmetric spaces," *Phys. Rept.* **394** (2004) 41–156, arXiv:cond-mat/0304363 [cond-mat].

[31] M. R. Zirnbauer, *The Oxford Handbook of Random Matrix Theory*, ch. Symmetry Classes, pp. 43–65. Oxford Univ. Press, 2011. arXiv:1001.0722 [math-ph].

[32] D. Bernard and A. LeClair, "A Classification of Non-Hermitian Random Matrices," in *Statistical Field Theories*, vol. 73 of *NATO Science Series*, pp. 207–214. 2002. arXiv:cond-mat/0110649 [cond-mat.dis-nn].

[33] B. Eynard, "Universal distribution of random matrix eigenvalues near the 'birth of a cut' transition," *J. Stas. Mech.* **2006** (2006) P07005, arXiv:math-ph/0605064 [math-ph].

[34] M. Bertola and S. Y. Lee, "First Colonization of a Spectral Outpost in Random Matrix Theory," *Const. Approx.* **30** (2009) 225–263, arXiv:0711.3625 [math-ph].

[35] E. Brézin and N. Deo, "Correlations and symmetry breaking in gapped matrix models," *Phys. Rev.* **E59** (1999) 3901–3910, arXiv:cond-mat/9805096 [cond-mat].

[36] B. Eynard, *Random Matrices, Random Processes and Integrable Systems*, ch. Formal Matrix Integrals and Combinatorics of Maps, pp. 415–442. CRM Series in Mathematical Physics. Springer New York, 2011. arXiv:math-ph/0611087 [math-ph].

[37] M. A. Bershadsky and A. A. Migdal, "Ising Model of a Randomly Triangulated Random Surface as a Definition of Fermionic String Theory," *Phys. Lett.* **B174** (1986) 393–398.

[38] V. A. Kazakov, "Ising model on a dynamical planar random lattice: Exact solution," *Phys. Lett.* **A119** (1986) 140–144.

[39] I. K. Kostov, "O(n) Vector Model on a Planar Random Lattice: Spectrum of Anomalous Dimensions," *Mod. Phys. Lett.* **A4** (1989) 217–226.

[40] M. Gaudin and I. Kostov, "O(N) Model on a Fluctuating Planar Lattice: Some Exact Results," *Phys. Lett.* **B220** (1989) 200–206.

[41] B. Eynard and C. Kristjansen, "Exact solution of the $O(n)$ model on a random lattice," *Nucl. Phys.* **B455** (1995) 577–618, arXiv:hep-th/9506193 [hep-th].

[42] B. Eynard and C. Kristjansen, "An iterative solution of the three-colour problem on a random lattice," *Nucl. Phys.* **B516** (1998) 529–542, arXiv:cond-mat/9710199 [cond-mat].

[43] I. K. Kostov, "Exact solution of the three color problem on a random lattice," *Phys. Lett.* **B549** (2002) 245–252, arXiv:hep-th/0005190 [hep-th].

[44] P. H. Ginsparg, "Matrix models of 2d gravity," arXiv:hep-th/9112013 [hep-th].

[45] I. K. Kostov, "Exact solution of the six vertex model on a random lattice," *Nucl. Phys.* **B575** (2000) 513–534, arXiv:hep-th/9911023 [hep-th].

[46] M. Bertola, "Boutroux curves with external field: equilibrium measures without a minimization problem," *Analysis and Mathematical Physics* **1** (2011) 167–211, arXiv:0705.3062 [nlin.SI].

[47] J. D. Fay, *Theta Functions on Riemann Surfaces*, vol. 352 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, 1973.

[48] B. Eynard and A. Prats Ferrer, "Topological expansion of the chain of matrices," *JHEP* **0907** (2009) 096, arXiv:0805.1368 [math-ph].

[49] A. A. Migdal, "Loop Equations and $1/N$ Expansion," *Phys. Rept.* **102** (1983) 199–290.

[50] G. Borot and A. Guionnet, "Asymptotic expansion of beta matrix models in the multi-cut regime," arXiv:1303.1045 [math-ph].

[51] B. Eynard and N. Orantin, "Invariants of algebraic curves and topological expansion," *Commun. Num. Theor. Phys.* **1** (2007) 347–452, arXiv:math-ph/0702045 [math-ph].

[52] B. Eynard, "A concise expression for the ODE's of orthogonal polynomials," arXiv:math-ph/0109018.

[53] P. Deift, *Orthogonal polynomials and random matrices: a Riemann–Hilbert approach*, vol. 3 of *Courant Lecture Notes*. American Mathematical Soc., 2000.

[54] M. Jimbo and T. Miwa, "Solitons and Infinite Dimensional Lie Algebras," *Publ. Res. Inst. Math. Sci. Kyoto* **19** (1983) 943–1001.

[55] K. Ueno and K. Takasaki, "Toda lattice hierarchy," *Adv. Stud. Pure Math.* **4** (1984) 1–95.

[56] O. Babelon, D. Bernard, and M. Talon, *Introduction to Classical Integrable Systems*. Cambridge Monographs on Mathematical Physics. Cambridge Univ. Press, 2003.

[57] B. Eynard and M. L. Mehta, "Matrices coupled in a chain: I. Eigenvalue correlations," *J. Phys.* **A31** (1998) 4449–4456, arXiv:cond-mat/9710230 [cond-mat].

[58] M. Bertola, B. Eynard, and J. Harnad, "Differential systems for biorthogonal polynomials appearing in 2-matrix models and the associated Riemann-Hilbert problem," *Commun. Math. Phys.* **243** (2003) 193–240, arXiv:nlin/0208002 [nlin.SI].

[59] M. Bergère and B. Eynard, "Mixed correlation function and spectral curve for the 2-matrix model," *J. Phys.* **A39** (2006) 15091–15134, arXiv:math-ph/0605010.

[60] C. Itzykson and J.-B. Zuber, "The Planar Approximation. II," *J. Math. Phys.* **21** (1980) 411–421.

[61] Harish-Chandra, "Differential Operators on a Semisimple Lie Algebra," *Amer. J. Math.* **79** (1957) 87–120.

[62] J. J. Duistermaat and G. J. Heckman, "On the Variation in the cohomology of the symplectic form of the reduced phase space," *Invent. Math.* **69** (1982) 259–268.

[63] A. Morozov, "Pair correlator in the Itzykson–Zuber integral," *Mod. Phys. Lett.* **A7** (1992) 3503–3508, arXiv:hep-th/9209074 [hep-th].

[64] B. Eynard and A. Prats Ferrer, "2-matrix versus complex matrix model, integrals over the unitary group as triangular integrals," *Commun. Math. Phys.* **264** (2006) 115–144, arXiv:hep-th/0502041 [hep-th].

[65] M. Bertola and A. Prats Ferrer, "Harish-Chandra integrals as nilpotent integrals," arXiv:0801.3452 [math.GR].

[66] M. Bergère and B. Eynard, "Some properties of angular integrals," *J. Phys.* **A42** (2009) 265201, arXiv:0805.4482 [math-ph].

[67] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*. Oxford University Press, 2nd ed., 1997.

[68] P. Desrosiers, "Duality in random matrix ensembles for all β," *Nucl. Phys.* **B817** (2009) 224–251, arXiv:0801.3438 [math-ph].

Index